



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification</b> <sup>6</sup> : <b>C12N 15/12, C07K 14/705, A61K 38/17, C12N 5/10, C12Q 1/37, C12N 9/72, 15/85</b>	<b>A3</b>	<b>(11) International Publication Number:</b> <b>WO 98/55508</b> <b>(43) International Publication Date:</b> 10 December 1998 (10.12.98)
<b>(21) International Application Number:</b> PCT/JP98/02445 <b>(22) International Filing Date:</b> 3 June 1998 (03.06.98)  <b>(30) Priority Data:</b> 9/144948 3 June 1997 (03.06.97) JP  <b>(71) Applicants (for all designated States except US):</b> SAGAMI CHEMICAL RESEARCH CENTER [JP/JP]; 4-1, Nishi-Ohnuma 4-chome, Sagamihara-shi, Kanagawa 229-0012 (JP). PROTEGENE INC. [JP/JP]; 2-20-3, Naka-cho, Meguro-ku, Tokyo 153-0065 (JP).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> KATO, Seishi [JP/JP]; 3-46-50, Wakamatsu, Sagamihara-shi, Kanagawa 229-0014 (JP). SEKINE, Shingo [JP/JP]; Remonzu 101, 2-8-15, Atago, Ageo-shi, Saitama 362-0034 (JP). YAMAGUCHI, Tomoko [JP/JP]; 5-13-11, Takasago, Katsushika-ku, Tokyo 125-0054 (JP).  <b>(74) Agents:</b> AOYAMA, Tamotsu et al.; Aoyama & Partners, IMP Building, 3-7, Shiromi 1-chome, Chuo-ku, Osaka-shi, Osaka 540-0001 (JP).		<b>(81) Designated States:</b> AU, CA, JP, MX, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>  <b>(88) Date of publication of the international search report:</b> 25 March 1999 (25.03.99)
<b>(54) Title:</b> HUMAN PROTEINS HAVING TRANSMEMBRANE DOMAINS AND DNAs ENCODING THESE PROTEINS  <b>(57) Abstract</b>  Proteins comprising any of the amino acid sequences of SEQ ID NOS: 1 to 18 and DNAs encoding said proteins and comprising any of the nucleotide sequences of SEQ ID NOS: 19 to 36 are provided.		

# INTERNATIONAL SEARCH REPORT

International Application No

PC1/JP 98/02445

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/705 A61K38/17 C12N5/10 C12Q1/37  
C12N9/72 C12N15/85

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KYTE J. ET AL.: "A SIMPLE METHOD FOR DISPLAYING THE HYDROPATHIC CHARACTER OF A PROTEIN" JOURNAL OF MOLECULAR BIOLOGY, vol. 157, no. 1, 5 May 1982, pages 105-132, XP000609692 cited in the application	
A	LIBERT F. ET AL.: "SELECTIVE AMPLIFICATION AND CLONING OF FOUR NEW MEMBERS OF THE G PROTEIN-COUPLED RECEPTOR FAMILY" SCIENCE, vol. 244, 5 May 1989, pages 569-572, XP002041588	

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

### \* Special categories of cited documents :

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*Z\* document member of the same patent family

Date of the actual completion of the international search

22 September 1998

Date of mailing of the international search report

26.01.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP 98/02445

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-6 : all partially (see subject 1, extra sheet)

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-6 all partially.

A protein comprising an aminoacid sequence as in Seq.ID:1, encoding DNA, as in Seq.ID:19 and 37, related expression vector and transformed eukaryotic cell.

2. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:2, 20 and 38.

3. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:3, 21 and 39.

4. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:4, 22 and 40.

5. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:5, 23 and 41.

6. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:6, 24 and 42.

7. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:7, 25 and 43.

8. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:8, 26 and 44.

9. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:9, 27 and 45.

10. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:10, 28 and 46.

11. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:11, 29 and 47.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

12. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:12, 30 and 48.

13. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:13, 31 and 49.

14. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:14, 32 and 50.

15. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:15, 33 and 51.

16. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:16, 34 and 52.

17. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:17, 35 and 53.

18. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:18, 36 and 54.

1/19

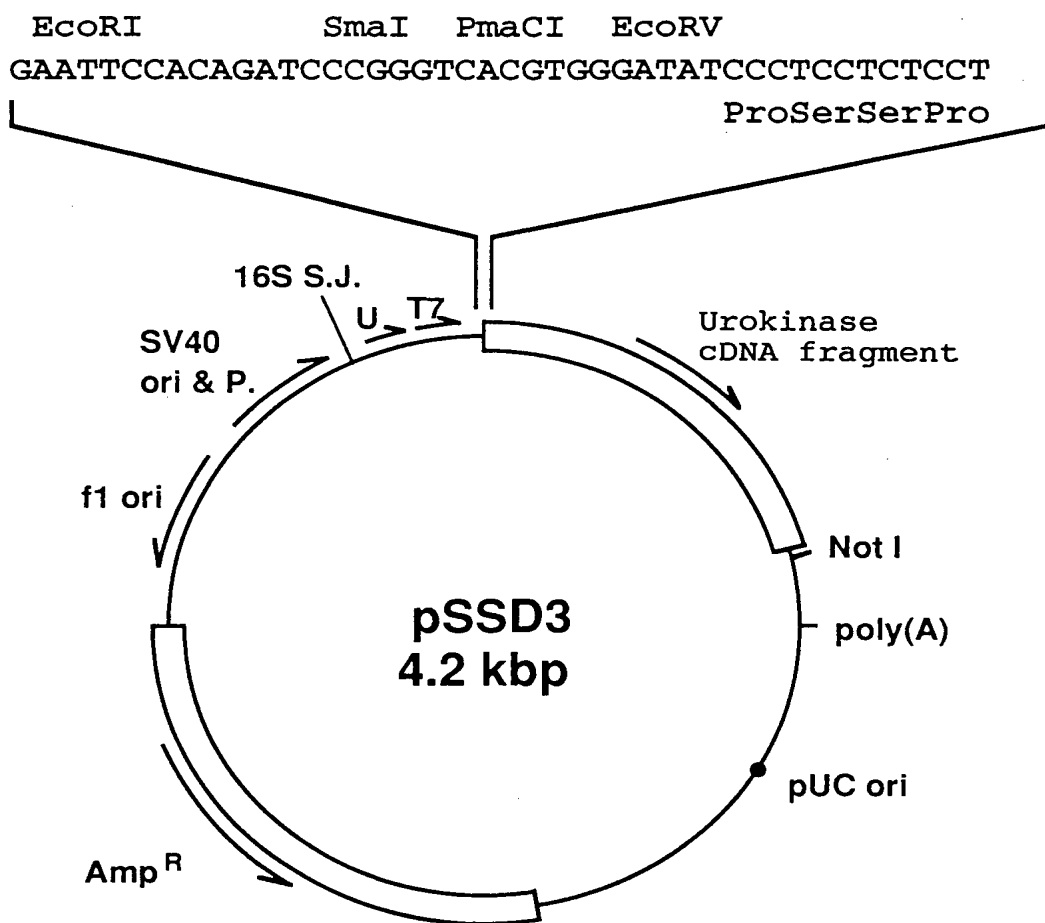


Fig.1

2/19

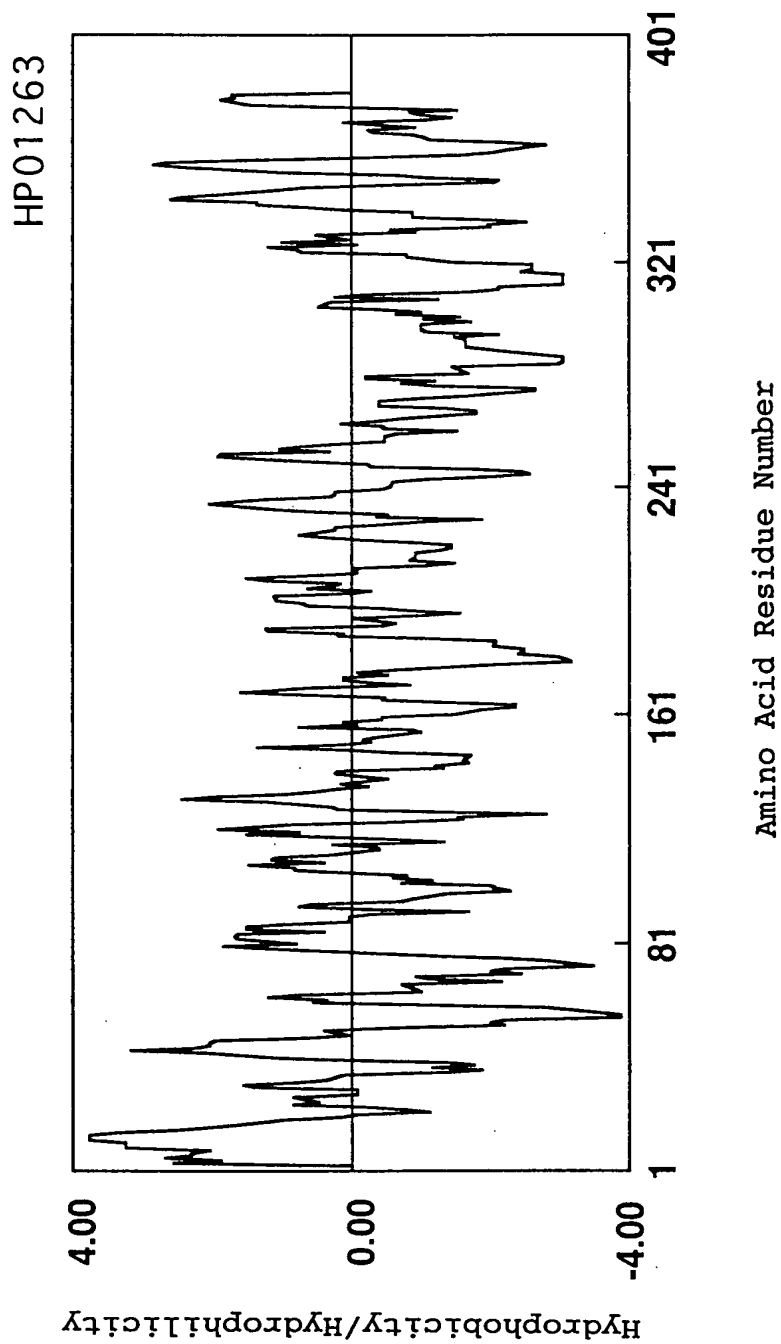
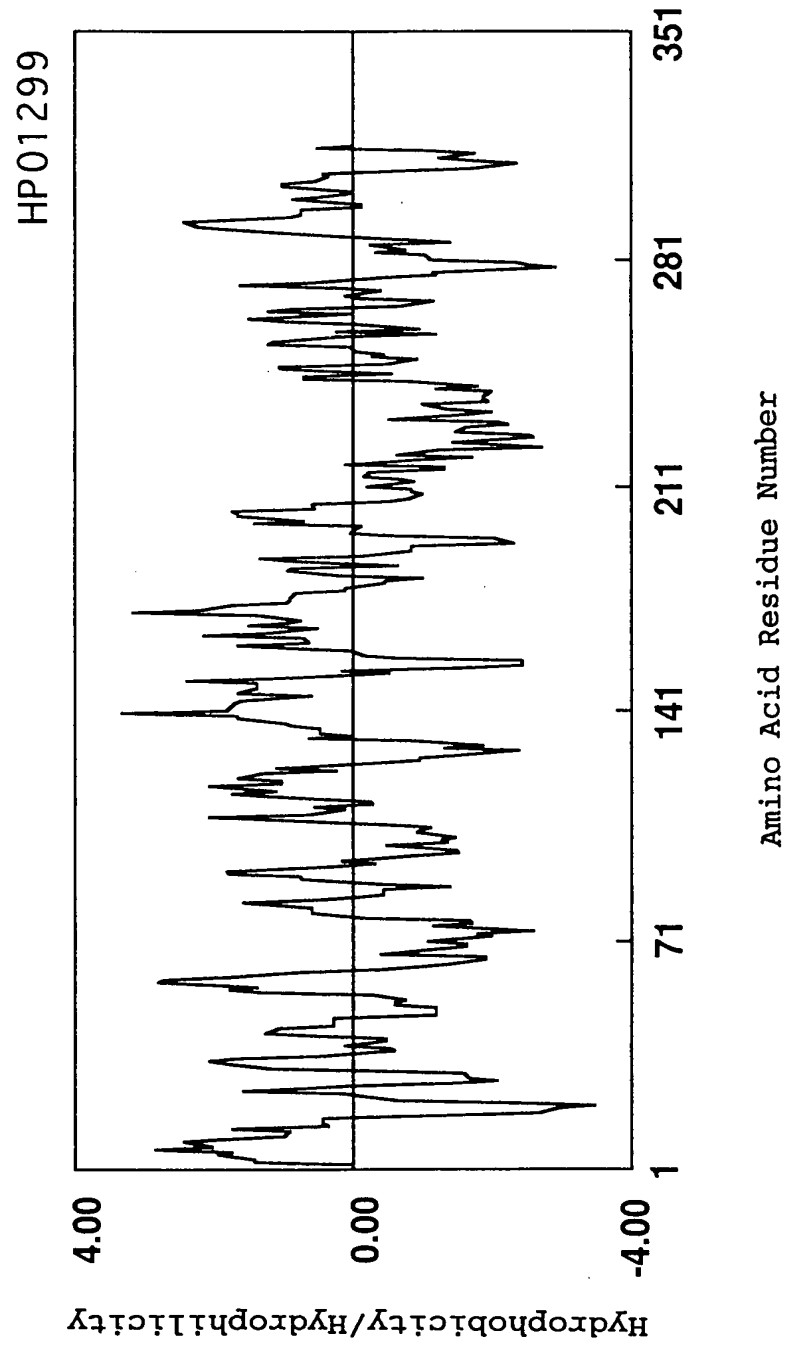


Fig.2



**Fig.3**



4/19

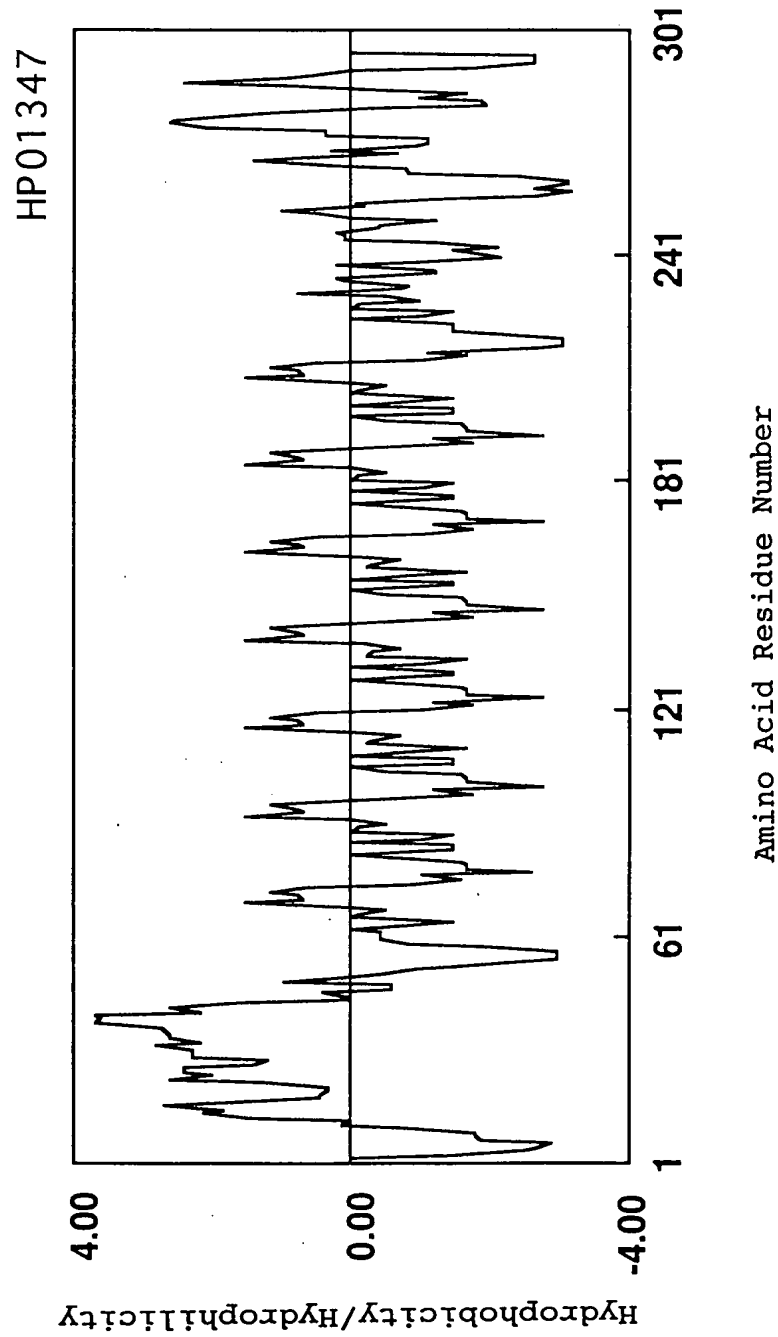


Fig.4

5/19

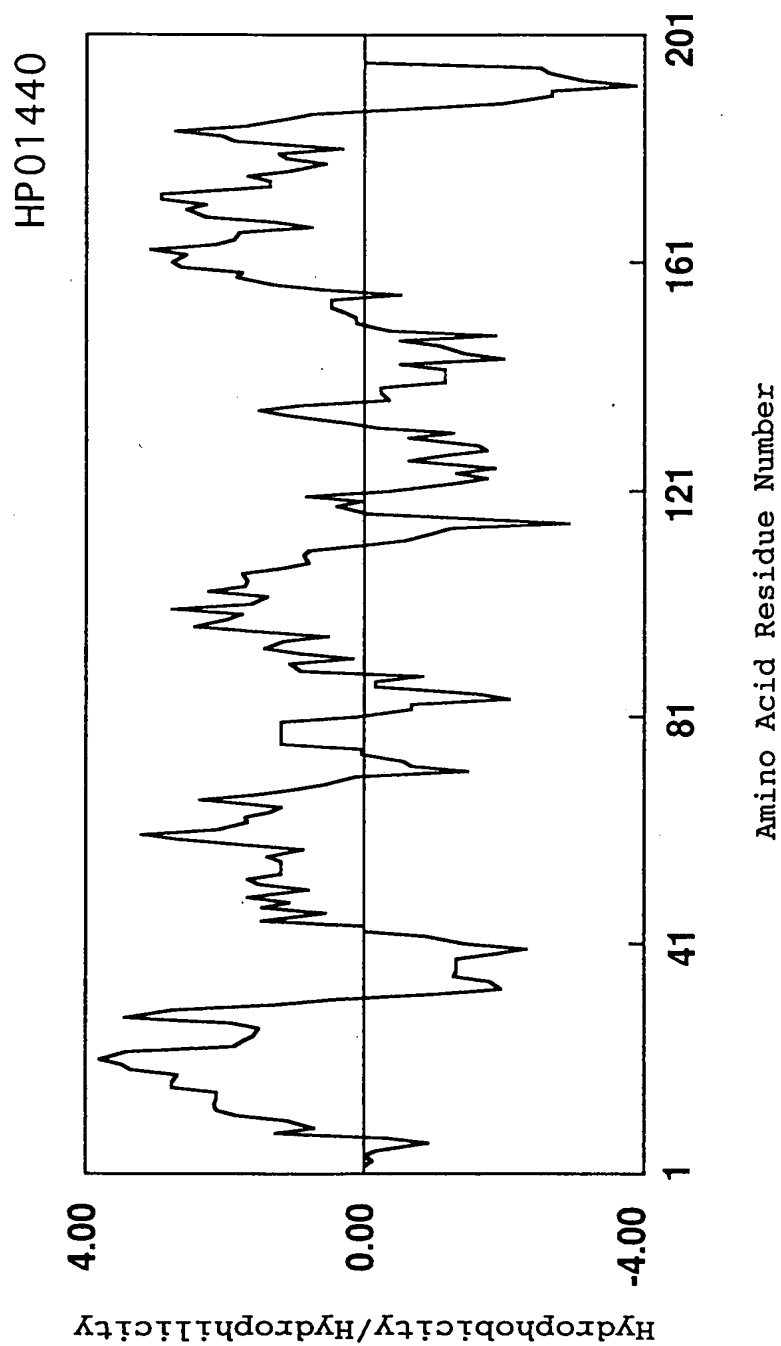


Fig.5

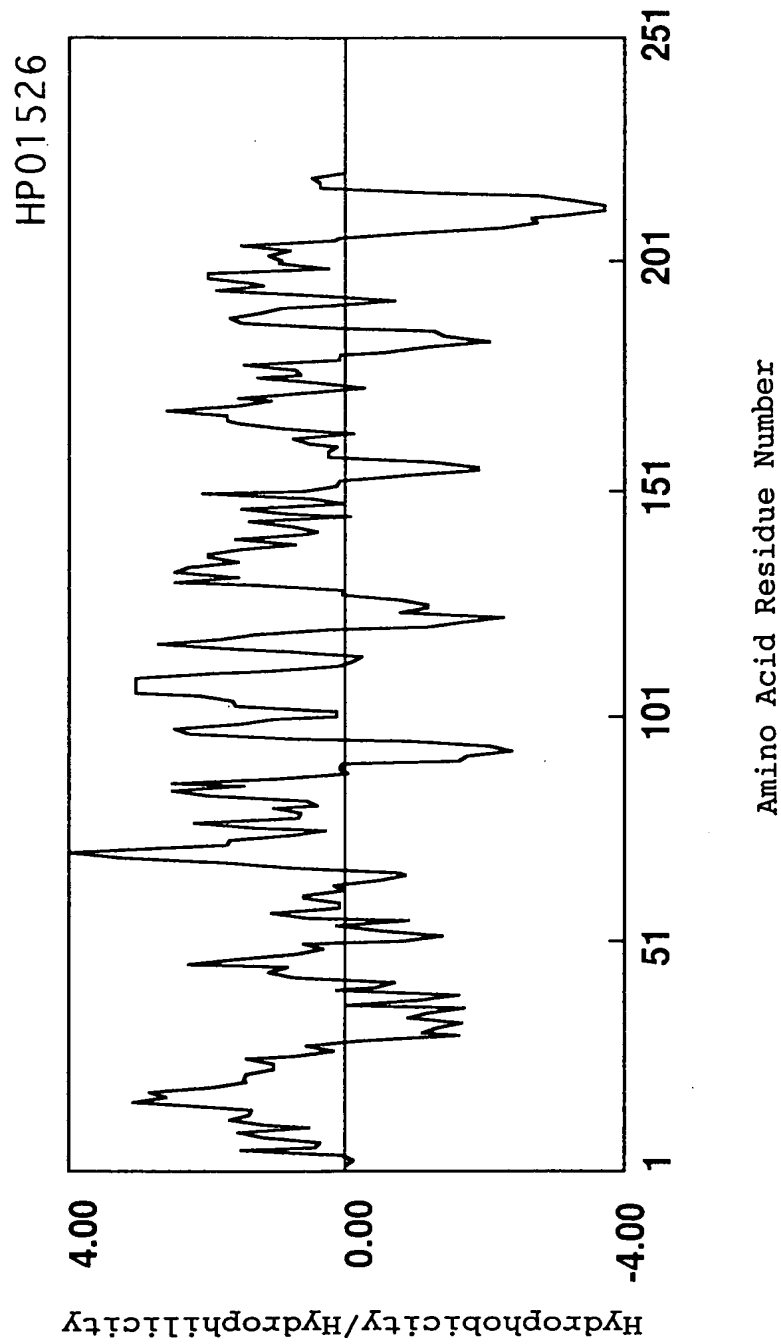
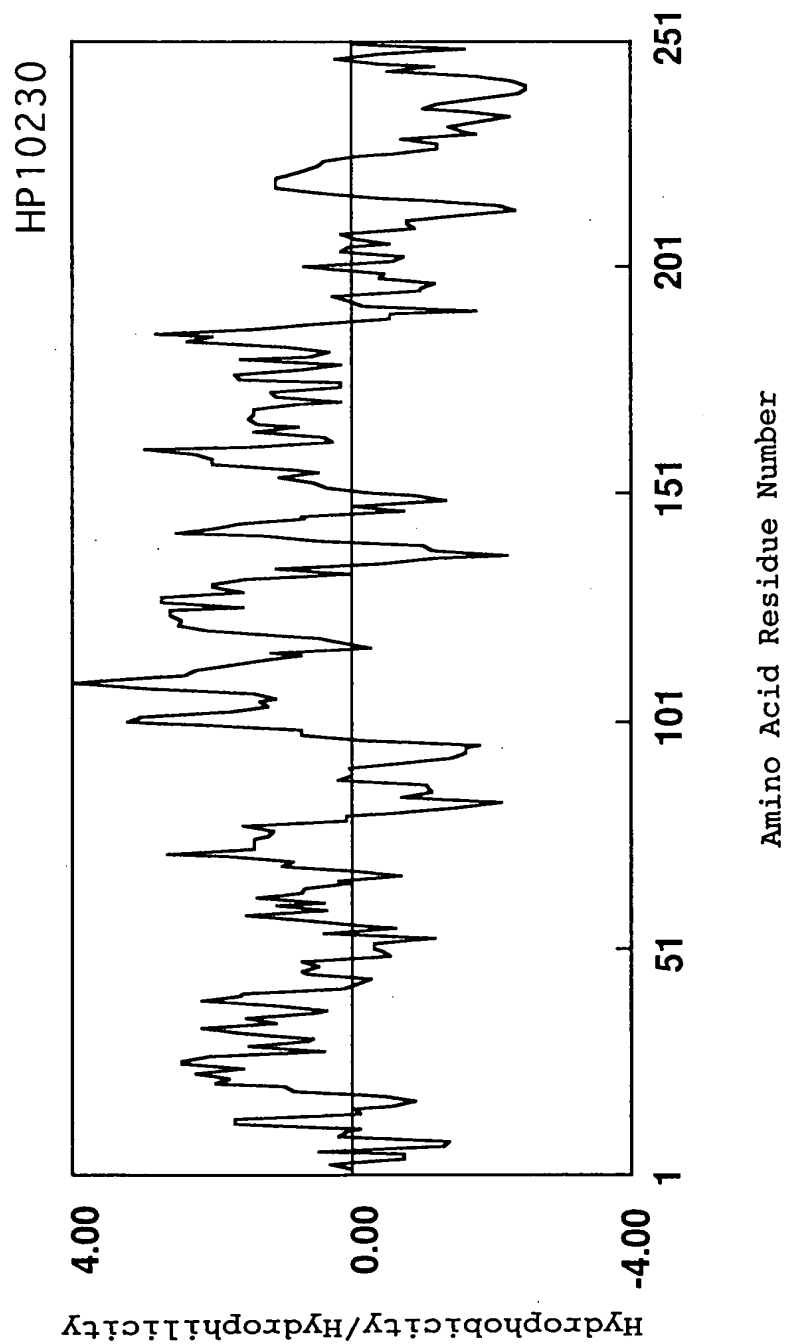


Fig.6



**Fig.7**

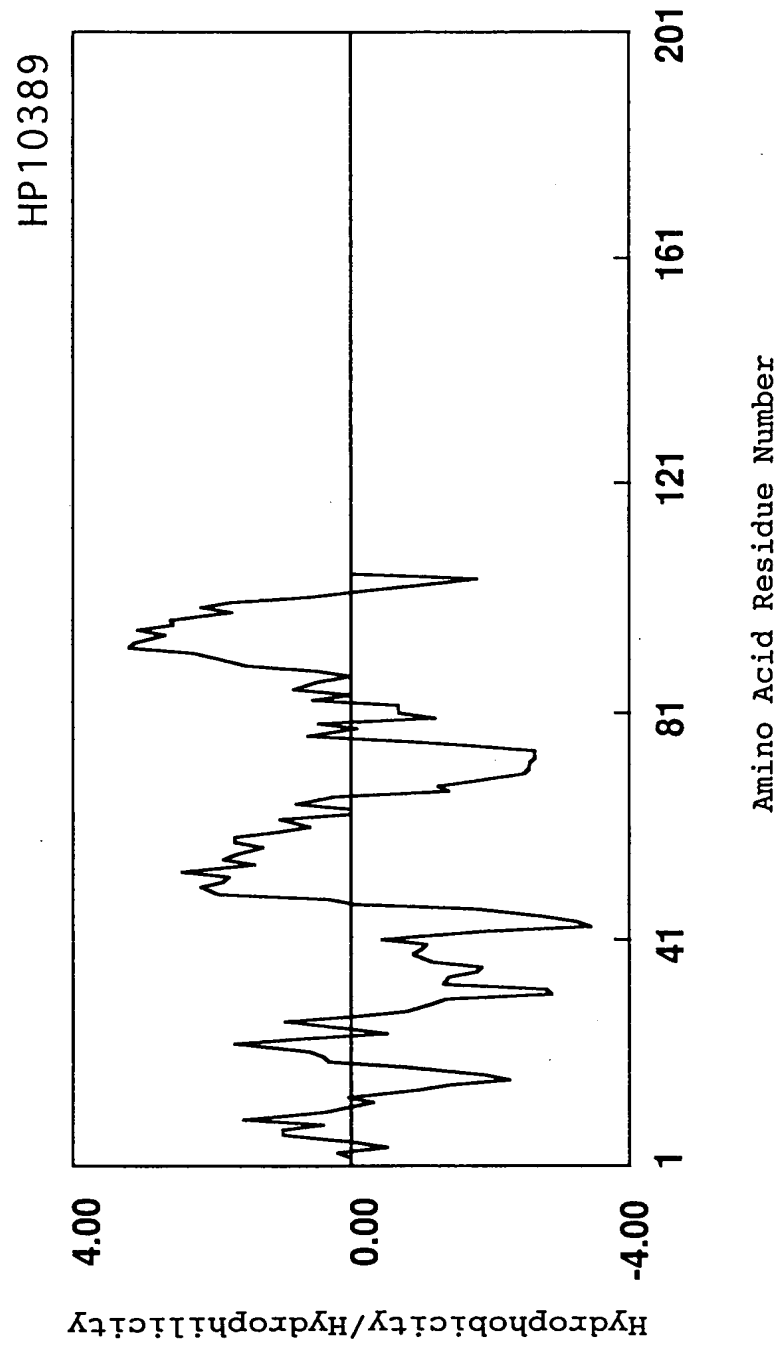


Fig.8

9/19

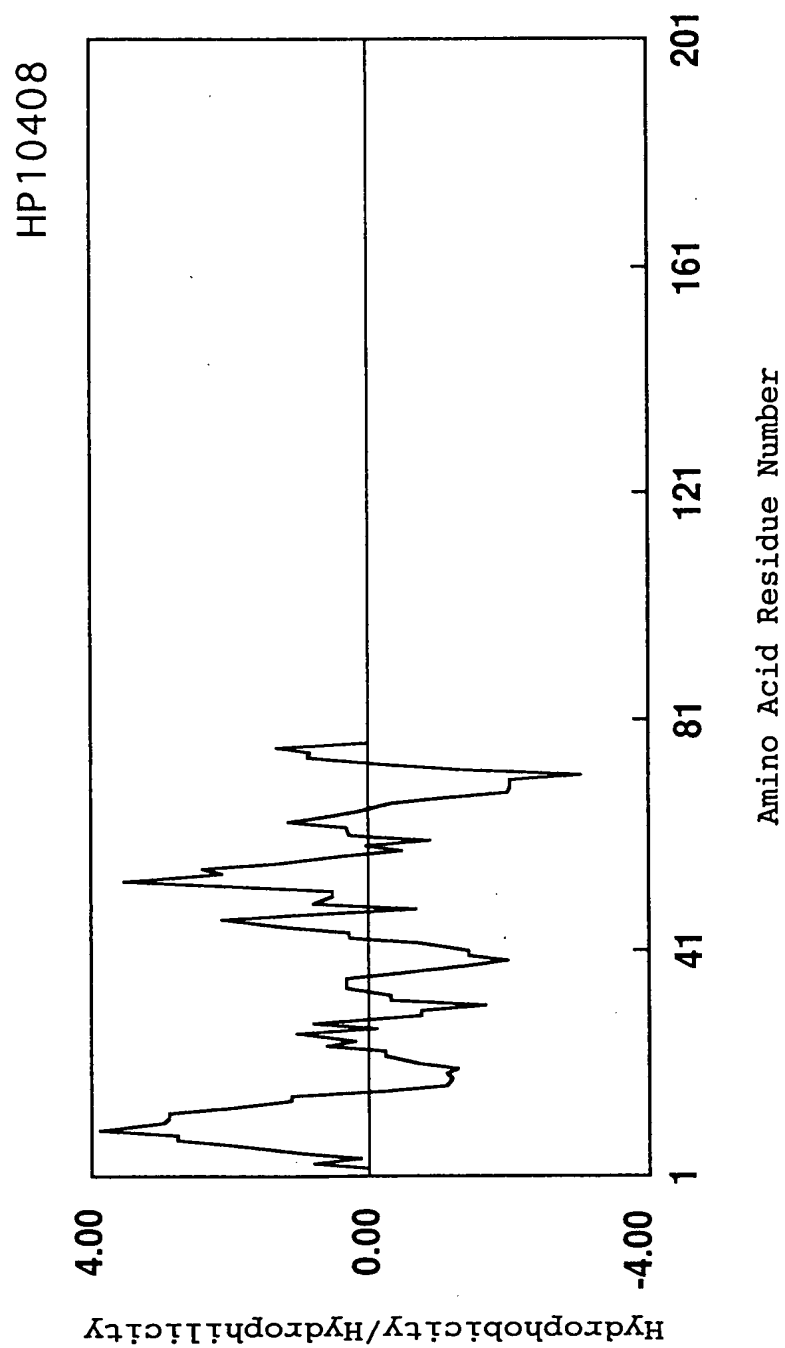


Fig.9

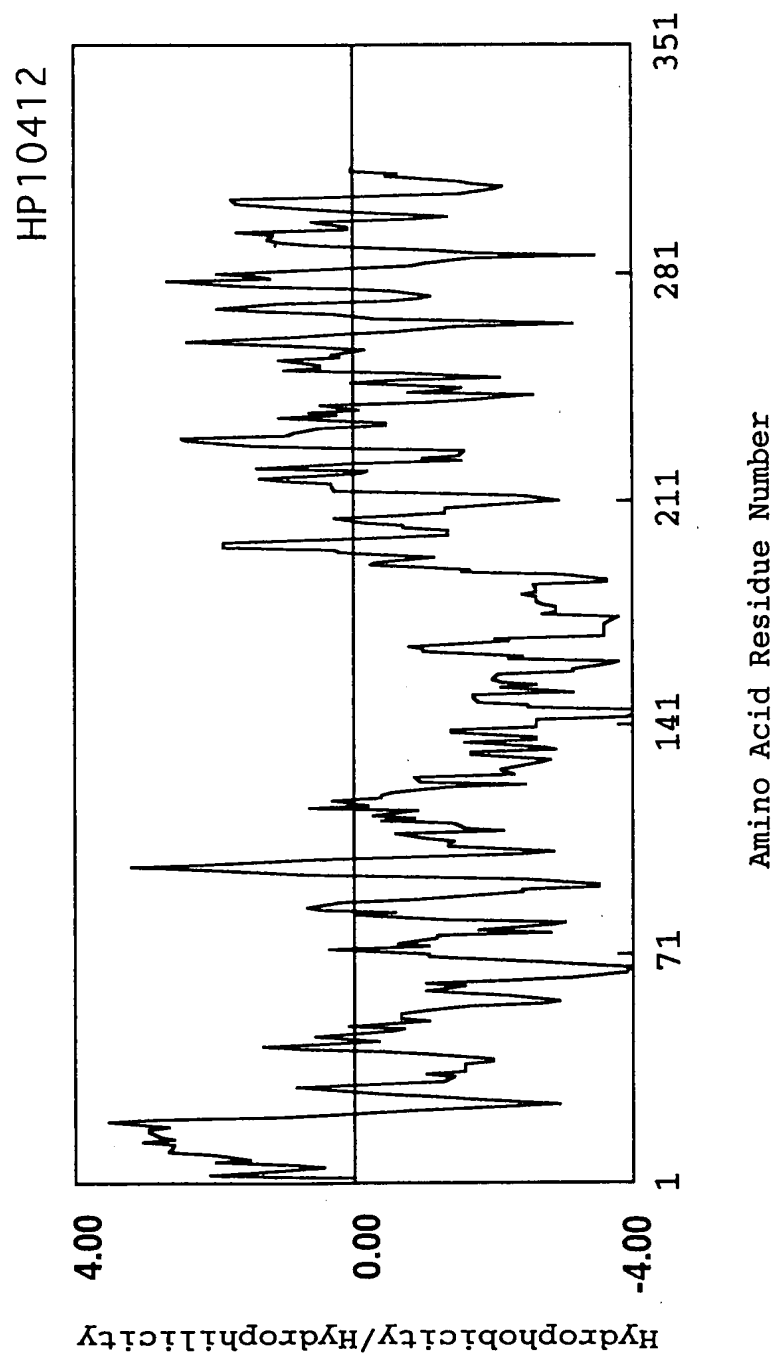


Fig.10

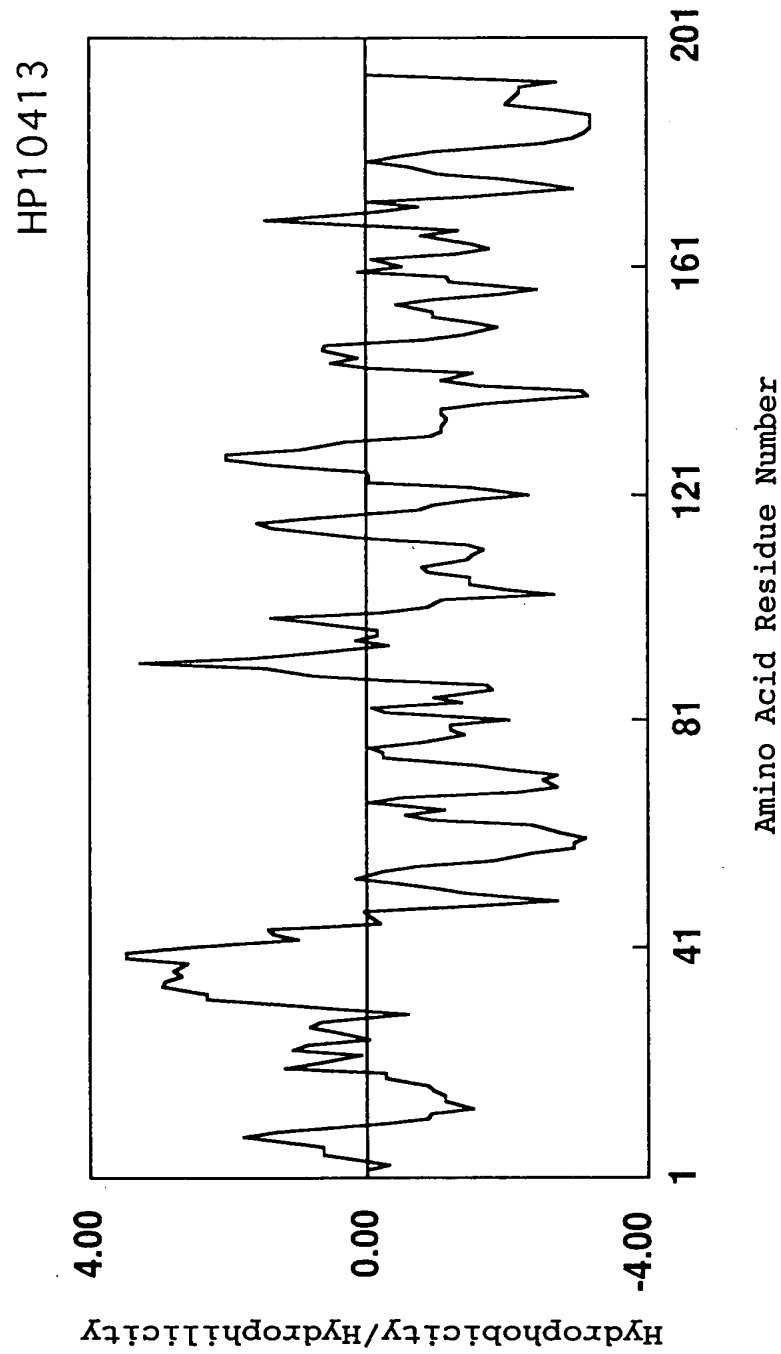


Fig.11



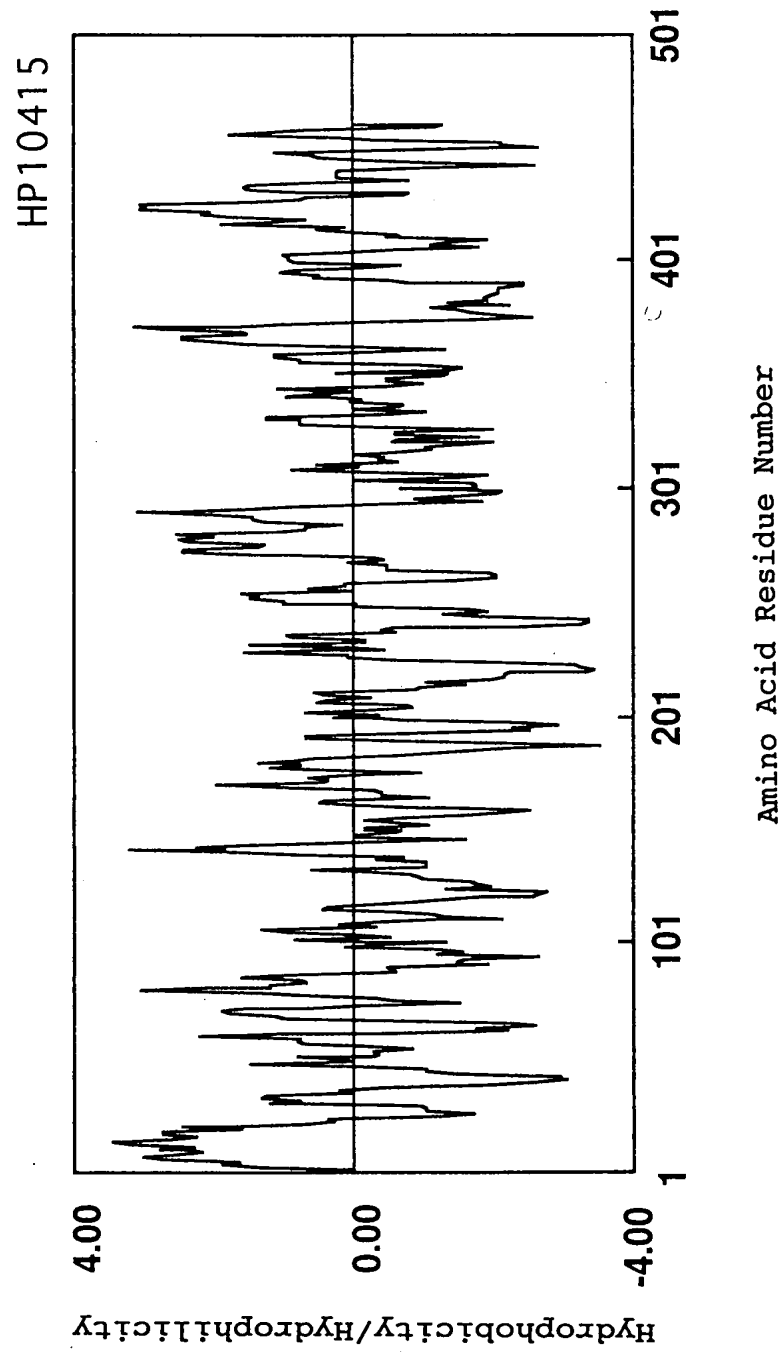


Fig.12

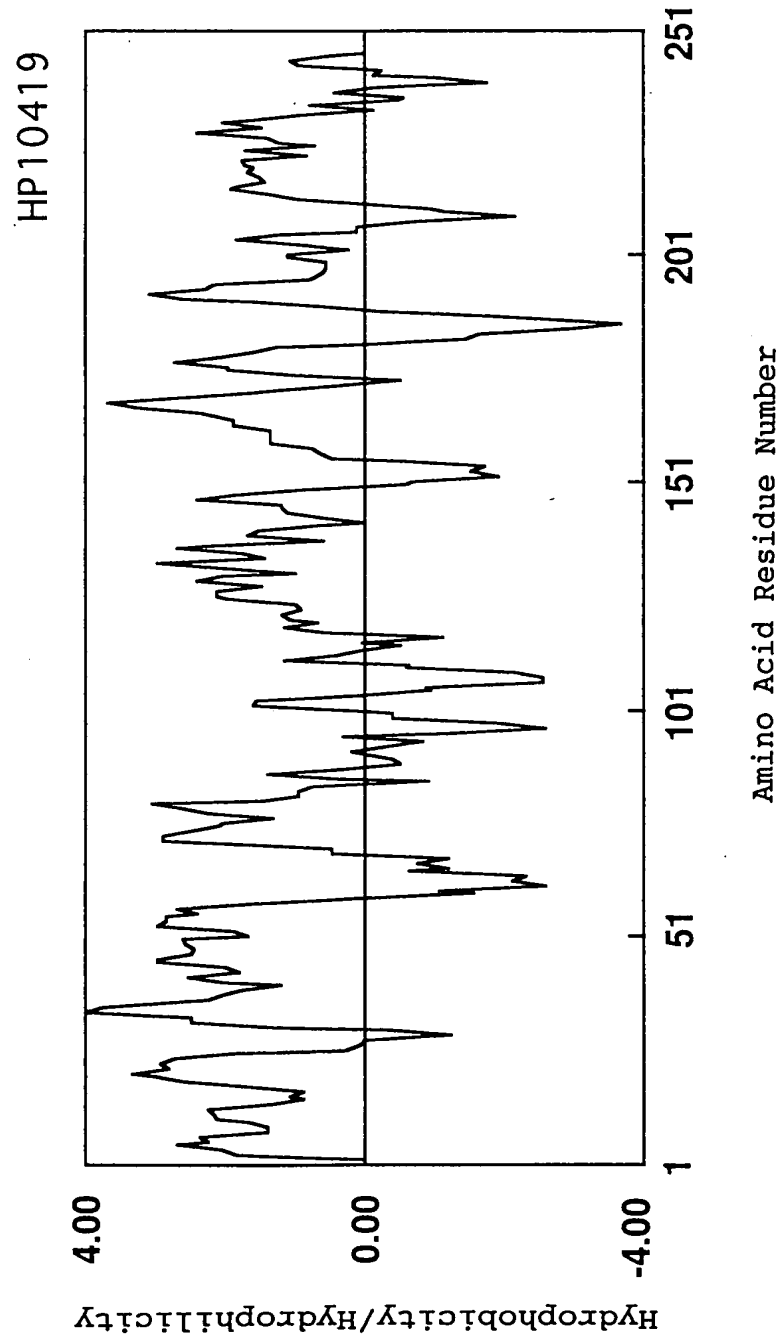


Fig.13

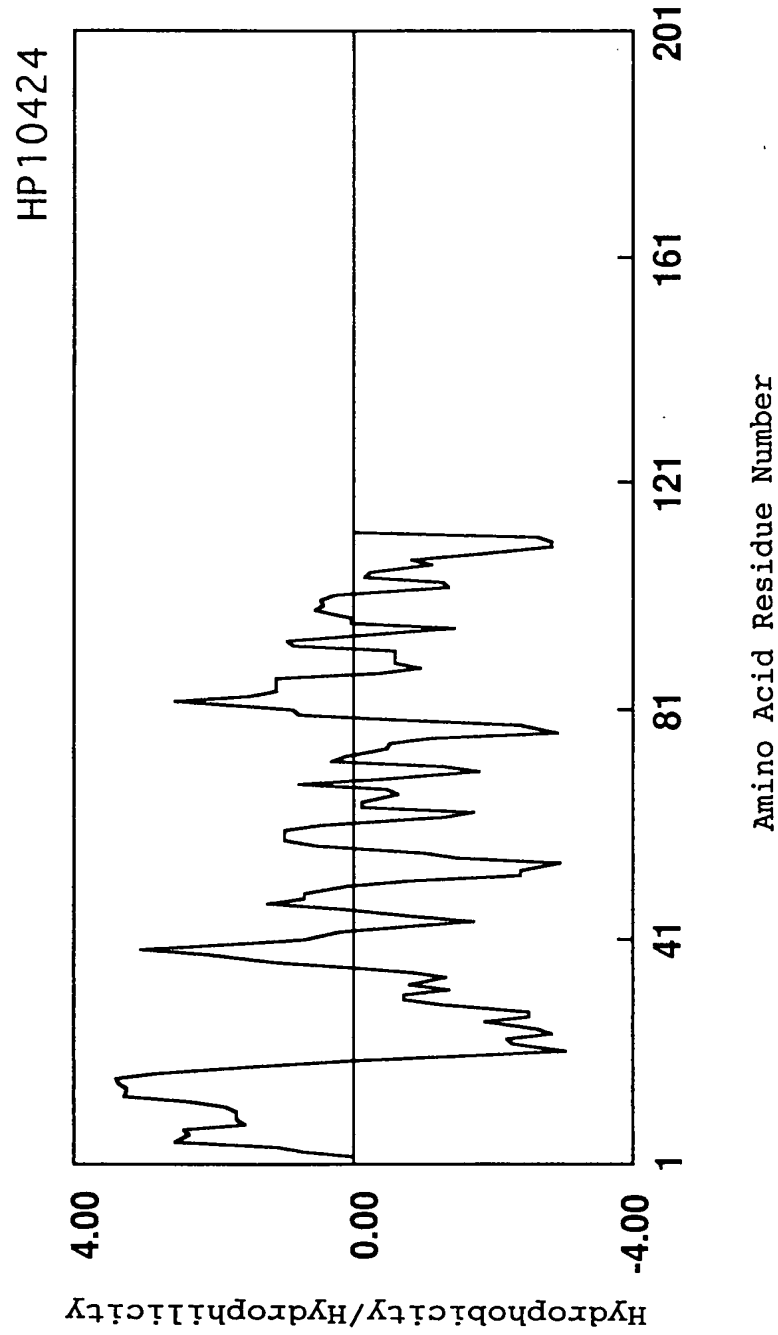


Fig.14

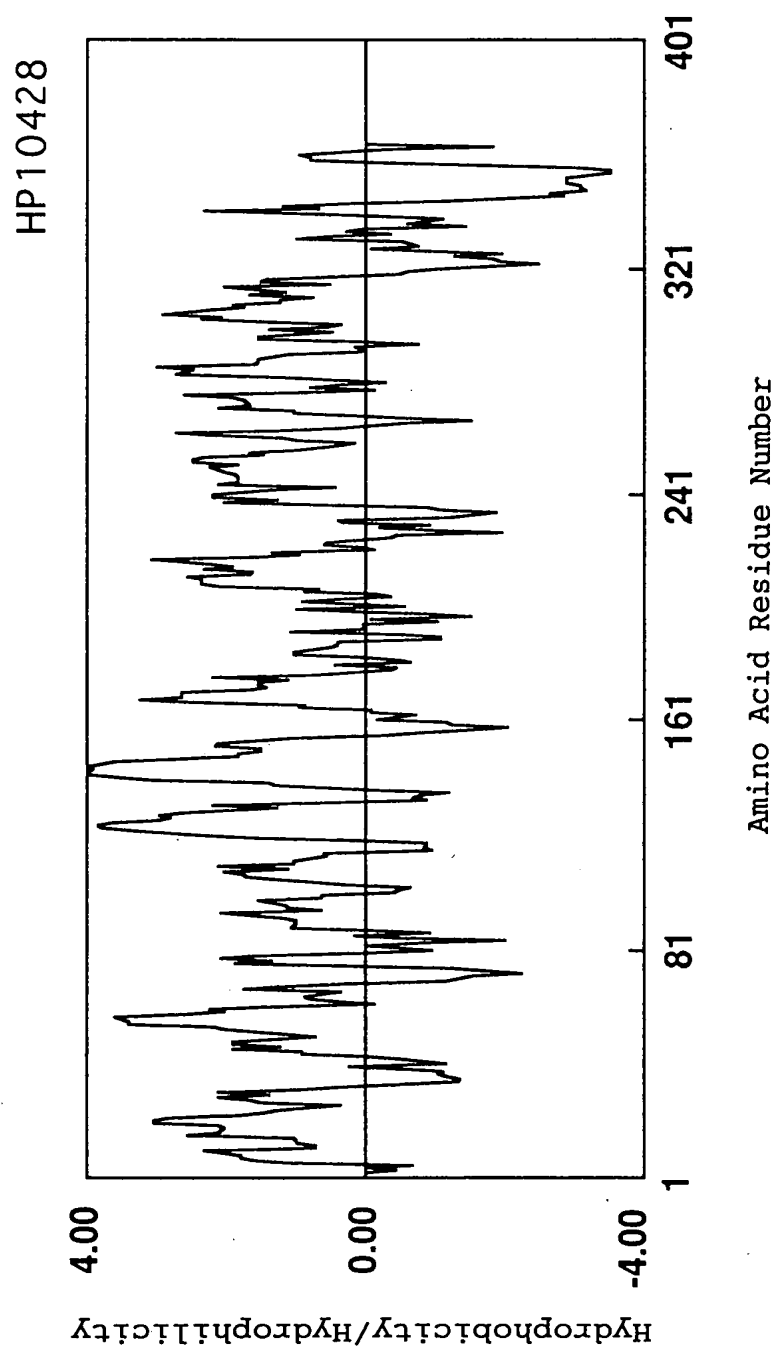


Fig.15

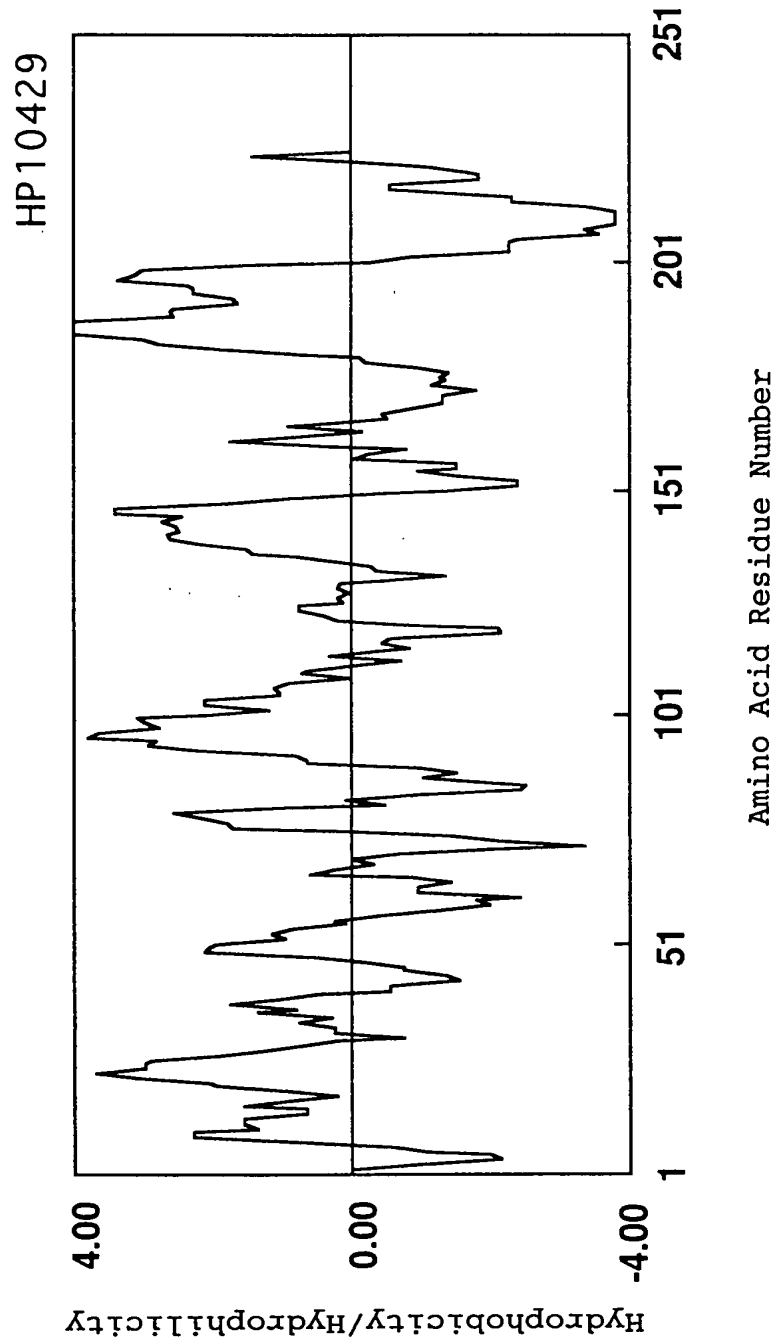
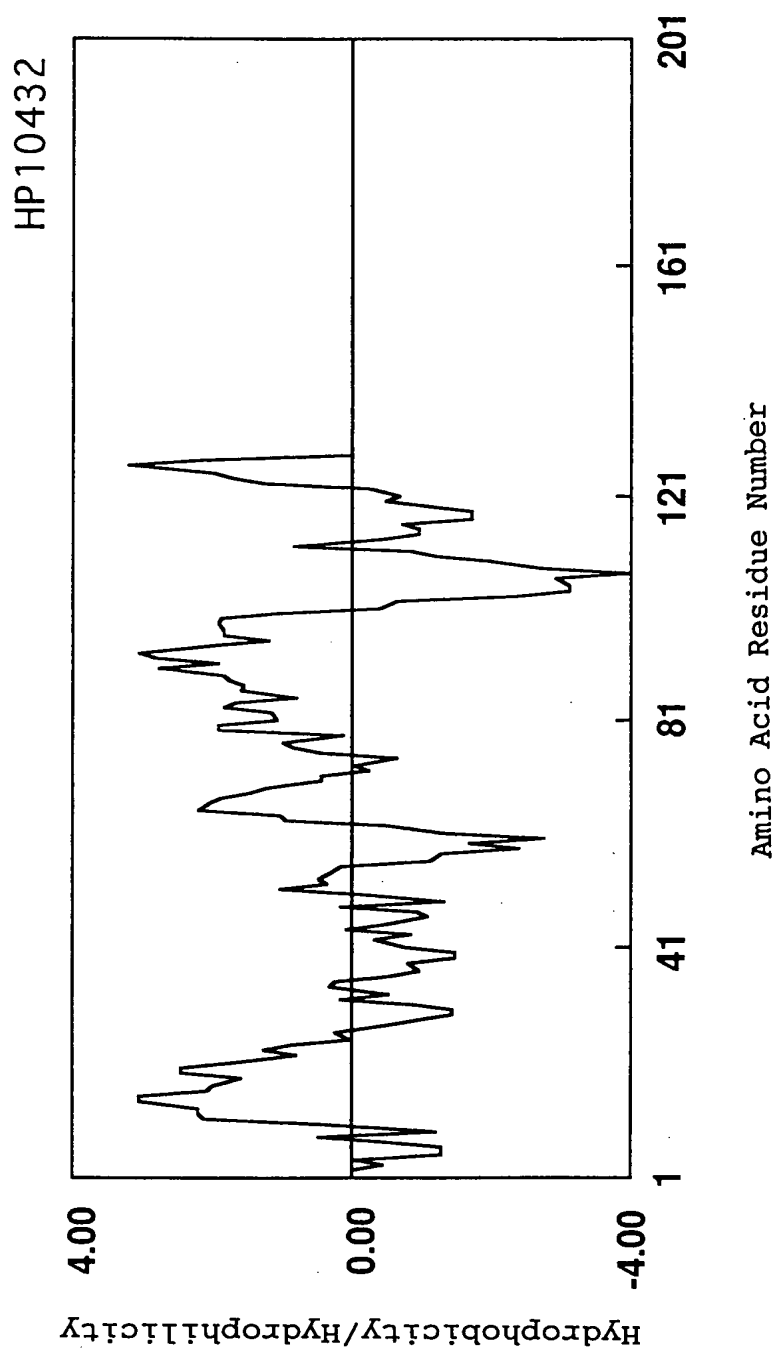


Fig.16



**Fig.17**

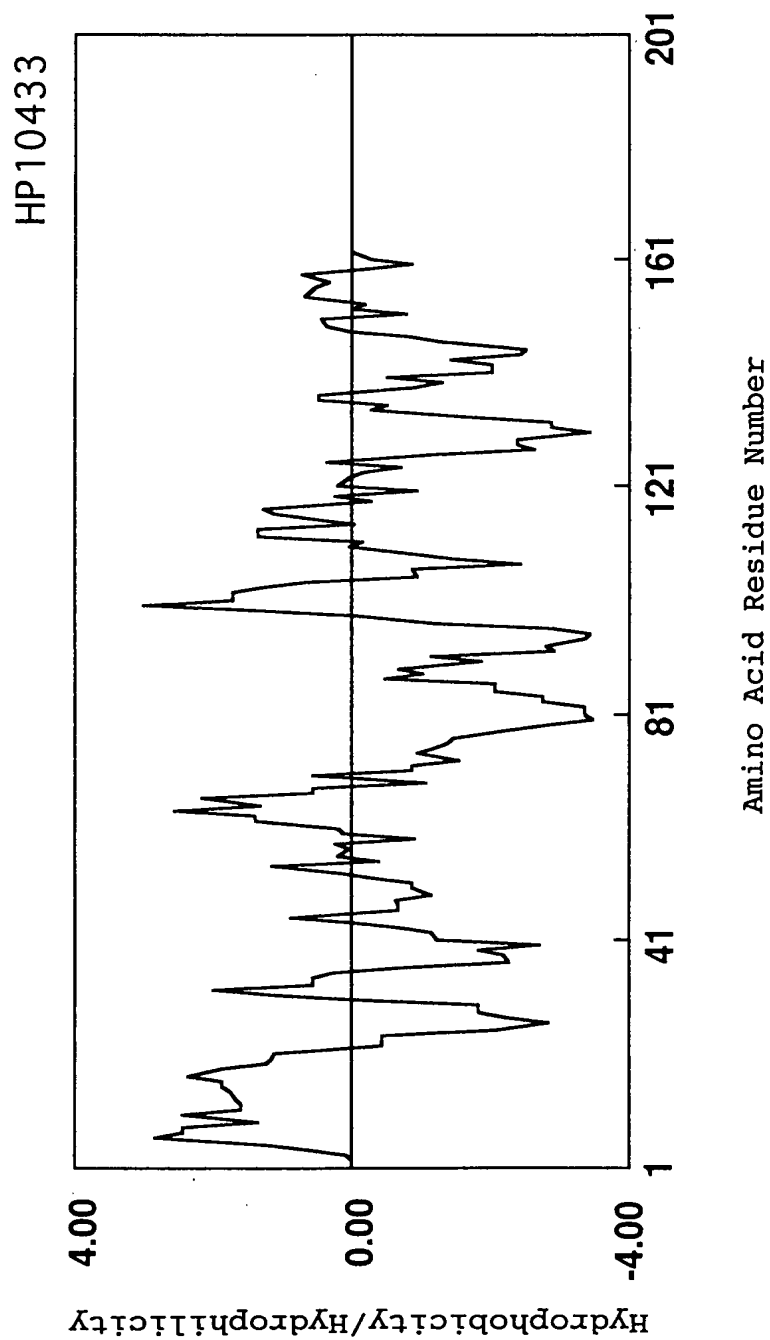


Fig.18

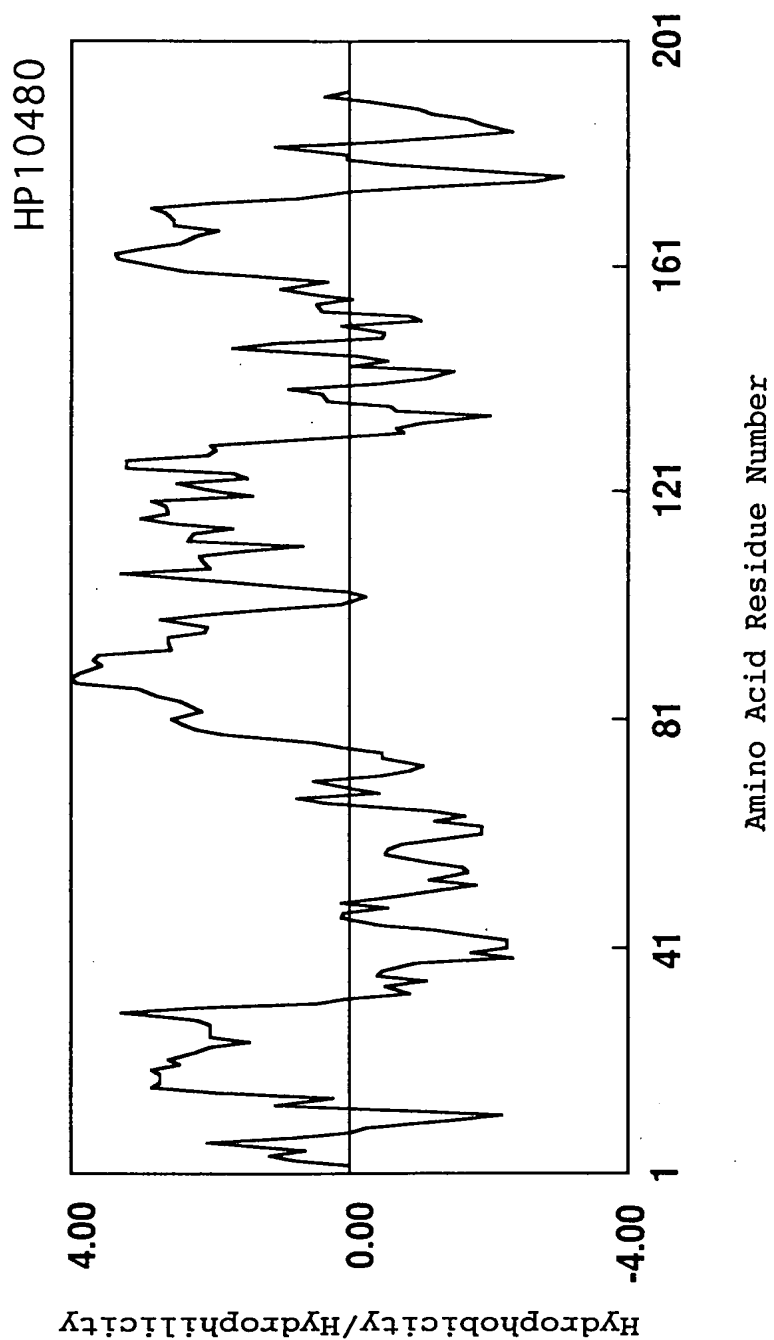


Fig.19



## DESCRIPTION

Human Proteins Having Transmembrane  
Domains and DNAs Encoding These Proteins

5

FIELD OF THE INVENTION

The present invention relates to human proteins having transmembrane domains and cDNAs encoding these proteins. The membrane proteins of this invention can be used as pharmaceuticals or as antigens for preparing antibodies against said proteins. The cDNAs of the invention can be used as probes for the gene diagnosis and gene sources for the gene therapy. The cDNAs can also be used as gene sources for large-scale production of the membrane proteins encoded by the same. The cells into which the genes encoding the membrane proteins are introduced for expression of such membrane proteins in large amounts can be used for detection of the corresponding ligands, screening of low molecular weight medicines, etc.

20 BACKGROUND OF THE INVENTION

Membrane proteins play important roles as signal receptors, ion channels, transporters, etc. for the material transportation or information transmission mediated by the cell membrane. For instance, they are known to serve as receptors for various cytokines, ion channels for sodium ion, potassium ion, chloride ion, etc., transporters for saccharides and amino acids, and so on. The genes for many of them have been cloned already.

In recent years, it was clarified that the abnormalities

of these membrane proteins are related to a number of hitherto cryptogenic diseases. For example, a gene for a membrane protein having 12 transmembrane domains was identified as the gene responsible for cystic fibrosis [Rommens, J. M. et al., Science 245: 1059-1065 (1989)]. It was also clarified that several membrane proteins act as the receptors when a virus infects the cells. For example, HIV-1 was revealed to infect into the cells through the mediation of a membrane protein fusin, a membrane protein on the T-cell membrane, having a CD-4 antigen and 7 transmembrane domains [Feng, Y. et al., Science 272: 872-877 (1996)]. Therefore, the discovery of new membrane proteins is anticipated to lead to the elucidation of the causes of many diseases, and the isolation of new genes coding for the membrane proteins is desired.

Heretofore, owing to the difficulty in their purification, many of membrane proteins have been isolated by an approach from the gene side. A general method is the so-called expression cloning which comprises transfection of a cDNA library in the animal cells to express the cDNA and detection of the cells expressing the target membrane protein on the membrane by an immunological technique using an antibody or a physiological technique for the change in the membrane permeability. However, this method is applicable only to cloning of a gene for a membrane protein with a known function.

In general, membrane proteins possess hydrophobic transmembrane domains inside the proteins which are synthesized in the ribosome. Said domains remain in the phospholipid to be trapped in the membrane. Accordingly, the evidence of the cDNA for encoding the membrane protein is provided by determination

of the whole base sequence of a full-length cDNA and detection of highly hydrophobic transmembrane domains in the amino acid sequence of the protein encoded by said cDNA.

As a result of the extensive study, there have successful-  
5 ly been obtained human proteins having transmembrane domains, particularly comprising any of the amino acid sequences of SEQ ID NOS: 1 to 18, by cloning cDNAs coding for proteins having transmembrane domains, particularly comprising any of the nucleotide sequences of SEQ ID NOS: 19 to 36, from a human  
10 full-length cDNA bank. The present invention is based on the above success.

#### SUMMARY OF THE INVENTION

A main object of the present invention is to provide novel  
15 human proteins having transmembrane domains, particularly comprising any of the amino acid sequences of SEQ ID NOS: 1 to 18. Another object of this invention is to provide DNAs coding for said novel proteins, particularly comprising any of the nucleotide sequences of SEQ ID NOS: 19 to 36. A further object  
20 of the invention is to provide expression vectors capable of in vitro translating said DNAs or expressing said DNAs in eukaryotic cells. A still further object of the invention is to provide transformed eukaryotic cells capable of expressing said DNAs to produce said proteins.

25 In one embodiment, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of the amino acid sequences of SEQ ID NOS: 1 to 18 and their fragments.

In another embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of the nucleotide sequences of SEQ ID NOS: 19 to 36.

5 In a further embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of the nucleotide sequences of SEQ ID NOS: 37 to 54.

#### 10 BRIEF DESCRIPTION OF DRAWINGS

Figure 1: A figure depicting the structure of the secretory signal sequence detection vector pSSD3.

Figure 2: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01263.

15 Figure 3: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01299.

Figure 4: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01347.

20 Figure 5: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01440.

Figure 6: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP01526.

Figure 7: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10230.

25 Figure 8: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10389.

Figure 9: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10408.

Figure 10: A figure depicting the hydrophobicity/hydro-

philicity profile of the protein encoded by clone HP10412.

Figure 11: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10413.

Figure 12: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10415.

Figure 13: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10419.

Figure 14: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10424.

Figure 15: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10428.

Figure 16: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10429.

Figure 17: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10432.

Figure 18: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10433.

Figure 19: A figure depicting the hydrophobicity/hydrophilicity profile of the protein encoded by clone HP10480.

#### BEST MODE FOR CARRING OUT INVENTION

The proteins of the present invention can be obtained, for example, by isolation from human organs, cell lines, etc., by chemical synthesis on the basis of the amino acid sequences as herein disclosed, or by recombinant DNA technology using the DNA encoding the transmembrane domains of the invention. Among them, adoption of the recombinant DNA technology is preferred. Specifically, each of the proteins may be prepared by in vitro transcription of a vector comprising the cDNA of the invention

to make RNA and in vitro translation using this RNA as a template to accomplish in vitro expression. Also, each of the proteins may be prepared in a large amount by the use of *Escherichia coli*, *Bacillus subtilis*, yeasts, animal cells, etc.

5 comprising a suitable expression vector having the DNA encoding such protein.

In the case of producing the protein of the invention by the use of a microorganism such as *Escherichia coli*, the translation region of the cDNA of the invention is constructed  
10 in an expression vector having an origin, a promoter, a ribosome-binding site, a cDNA-cloning site, a terminator, etc. that can be replicated in the microorganism and, after transformation of the host cells with said expression vector, the resultant transformant is incubated, whereby the protein  
15 encoded by said cDNA can be produced in a large amount in the microorganism. In that case, a protein fragment containing an optional region can be obtained by performing the expression with inserting an initiation codon and a termination codon before and after the optional translation region. Alternative-  
20 ly, a fusion protein with another protein can be expressed. Only a protein portion encoding said cDNA can be obtained by cleavage of said fusion protein with an appropriate protease.

For production of the protein of the invention by expression of DNA coding for such protein in eukaryotic cells,  
25 the translation region of said cDNA may be recombined into an expression vector for eukaryotic cells having a promoter, a splicing domain, a poly(A) addition site, etc., followed by introduction into eukaryotic cells so that the protein of the invention is produced as a membrane protein on the cell

membrane surface. Examples of the expression vector are pKA1, pED6\_dpc2, pCDM8, pSVK3, pMSG, pSVL, pBK-CMV, pBK-RSV, EBV vector, pRS, pYES2, etc. As the eukaryotic cells, there are exemplified mammalian animal culture cells (e.g. simian kidney cells COS7, chinese hamster ovary cells CHO), budding yeasts, Schizosaccharomyces pombe, silkworm cells, Xenopus laevis egg cells, etc., but any other eukaryotic cells may also be used insofar as the protein of the invention can be expressed on the membrane surface. In order to introduce the expression vector into eukaryotic cells, there may be adopted any conventional procedure such as electroporation, calcium phosphate method, liposome method or DEAE dextran method.

The proteins of the present invention include peptide fragments (5 or more amino acid residues) containing any partial amino acid sequence of the amino acid sequences of SEQ ID NOS: 1 to 18. These fragments can be used as antigens for preparation of the antibodies. Also, the proteins of the invention that have signal sequences appear in the form of maturation proteins on the cell surface, after the signal sequences are removed. Therefore, these maturation proteins shall come within the scope of the present invention. The N-terminal amino acid sequences of the maturation proteins can be easily identified by using the method for the cleavage-site determination in a signal sequence [Japan Patent Kokai No. 187100/96]. Further, many membrane proteins are subjected to the processing on the cell surface to be converted to the secretor forms. These secretor proteins or peptides shall come within the scope of the present invention. When glycosylation sites are present in the amino acid sequences, expression in

appropriate animal cells affords glycosylated proteins. Therefore, these glycosylated proteins or peptides also shall come within the scope of the invention.

The DNAs of the invention include all DNAs encoding the  
5 above-mentioned proteins. Said DNAs can be obtained using the method by chemical synthesis, the method by cDNA cloning, and so on.

Each of the cDNAs of the invention can be cloned from, for example, the cDNA libraries of the human cell origin. The cDNA  
10 is synthesized using as a template a poly(A)<sup>+</sup> RNA extracted from human cells. The human cells may be cells delivered from the human body, for example, by the operation or may be the culture cells. The cDNA can be synthesized by using any method selected from the Okayama-Berg method [Okayama, H. and Berg,  
15 P., Mol. Cell. Biol. 2: 161-170 (1982)], the Gubler-Hoffman method [Gubler, U. and Hoffman, J. Gene 25: 263-269 (1983)], and so on, but it is preferred to use the capping method [Kato, S. et al., Gene 150: 243-250 (1994)] as illustrated in Examples in order to obtain a full-length clone in an effective manner.

20 The primary selection of a cDNA encoding a human protein having transmembrane domains is performed by the sequencing of a partial base sequence of the cDNA clone selected at random from the cDNA libraries, sequencing of the amino acid sequence encoded by the base sequence, and recognition of the presence  
25 or absence of hydrophobic site(s) in the resulting N-terminal amino acid sequence region. Next, the secondary selection is carried out by determination of the whole base sequence by the sequencing and the protein expression by the in vitro translation. The ascertainment of the cDNA of the present



invention for encoding the protein having the secretory signal sequence is performed by using the signal sequence detection method [Yokoyama-Kobayashi, M. et al., Gene 163: 193-196 (1995)]. In other words, the ascertainment for the coding  
5 portion of the inserted cDNA fragment to function as a signal sequence is provided by fusing a cDNA fragment encoding the N-terminus of the target protein with a cDNA encoding the protease domain of urokinase and then expressing the resulting cDNA in COS7 cells to detect the urokinase activity in the cell  
10 culture medium. On the other hand, the N-terminal region is judged to remain in the membrane in the case where the urokinase activity is not detected in the cell culture medium.

The cDNAs of the invention are characterized by containing any of the nucleotide sequences of SEQ ID NOS: 19 to 36 or any  
15 of the nucleotide sequences of SEQ ID NOS: 37 to 54. Table 1 summarizes the clone number (HP number), the cells affording the cDNA, the total nucleotide number of the cDNA, and the number of the amino acid residues of the encoded protein, for each of the cDNAs.

Table 1

Sequence Number	HP Number	Cells	Number of Nucleotides	Number of Amino Acid Residues
1, 19, 37	HP01263	Liver	1502	382
2, 20, 38	HP01299	Liver	1349	317
3, 21, 39	HP01347	Liver	1643	296
4, 22, 40	HP01440	Stomach cancer	729	197
5, 23, 41	HP01526	Stomach cancer	1322	221
6, 24, 42	HP10230	Stomach cancer	3045	251
7, 25, 43	HP10389	KB	653	106
8, 26, 44	HP10408	Stomach cancer	439	78
9, 27, 45	HP10412	Stomach cancer	1131	314
10, 28, 46	HP10413	Stomach cancer	1875	195
11, 29, 47	HP10415	Stomach cancer	1563	462
12, 30, 48	HP10419	Stomach cancer	2030	247
13, 31, 49	HP10424	Stomach cancer	493	113
14, 32, 50	HP10428	KB	2044	365
15, 33, 51	HP10429	Stomach cancer	1043	226
16, 34, 52	HP10432	Liver	972	129
17, 35, 53	HP10433	Liver	695	163
18, 36, 54	HP10480	Stomach cancer	1914	193

Hereupon, the same clone as any of the cDNAs of the invention can be easily obtained by screening of the cDNA libraries constructed from the cell line or the human tissues employed in the invention, by the use of an oligonucleotide probe synthesized on the basis of the corresponding cDNA nucleotide sequence of SEQ ID NOS: 37 to 54.

In general, the polymorphism due to the individual difference is frequently observed in human genes. Therefore, any cDNA that is subjected to insertion or deletion of one or plural nucleotides and/or substitution with other nucleotides

in SEQ ID NOS: 37 to 54 shall come within the scope of the invention.

In a similar manner, any protein that is produced by these modifications comprising insertion or deletion of one or plural nucleotides and/or substitution with other nucleotides shall come within the scope of the present invention, as far as said protein possesses the activity of the corresponding protein having the amino acid sequence of SEQ ID NOS: 1 to 18.

The cDNAs of the invention include cDNA fragments (more than 10 bp) containing any partial nucleotide sequence of the nucleotide sequence of SEQ ID NOS: 19 to 36 or of the nucleotide sequence of SEQ ID NOS: 37 to 54. Also, DNA fragments consisting of a sense chain and an anti-sense chain shall come within this scope. These DNA fragments can be used as the probes for the gene diagnosis.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate

genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

- 5           Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave
- 10 the mRNA transcribed from the gene (Albert and Morris, 1994, Trends Pharmacol. Sci. 15(7): 250-254; Lavarosky et al., 1997, Biochem. Mol. Med. 62(1): 11-22; and Hampel, 1998, Prog. Nucleic Acid Res. Mol. Biol. 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that
- 15 have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified
- 20 genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to
- 25 the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through

insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, Bioessays 14(9): 629-633; Zwaal et al., 1993, Proc. Natl. Acad. Sci. USA 90(16): 7431-7435; Clark et al., 1994, Proc. Natl. Acad. Sci. USA 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour et al., 1988, Nature 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at

least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences

complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably  
5 highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent  
10 conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

0944553 8525460

Table 2

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>‡</sup>	Hybridization Temperature and Buffer <sup>†</sup>	Wash Temperature and Buffer <sup>†</sup>
A	DNA : DNA	≥50	65°C; 1×SSC -or- 42°C; 1×SSC,50% formamide	65°C; 0.3×SSC
B	DNA : DNA	<50	T <sub>B</sub> *; 1×SSC	T <sub>B</sub> *; 1×SSC
C	DNA : RNA	≥50	67°C; 1×SSC -or- 45°C; 1×SSC,50% formamide	67°C; 0.3×SSC
D	DNA : RNA	<50	T <sub>D</sub> *; 1×SSC	T <sub>D</sub> *; 1×SSC
E	RNA : RNA	≥50	70°C; 1×SSC -or- 50°C; 1×SSC,50% formamide	70°C; 0.3×SSC
F	RNA : RNA	<50	T <sub>F</sub> *; 1×SSC	T <sub>F</sub> *; 1×SSC
G	DNA : DNA	≥50	65°C; 4×SSC -or- 42°C; 4×SSC,50% formamide	65°C; 1×SSC
H	DNA : DNA	<50	T <sub>H</sub> *; 4×SSC	T <sub>H</sub> *; 4×SSC
I	DNA : RNA	≥50	67°C; 4×SSC -or- 45°C; 4×SSC,50% formamide	67°C; 1×SSC
J	DNA : RNA	<50	T <sub>J</sub> *; 4×SSC	T <sub>J</sub> *; 4×SSC
K	RNA : RNA	≥50	70°C; 4×SSC -or- 50°C; 4×SSC,50% formamide	67°C; 1×SSC
L	RNA : RNA	<50	T <sub>L</sub> *; 2×SSC	T <sub>L</sub> *; 2×SSC
M	DNA : DNA	≥50	50°C; 4×SSC -or- 40°C; 6×SSC,50% formamide	50°C; 2×SSC
N	DNA : DNA	<50	T <sub>N</sub> *; 6×SSC	T <sub>N</sub> *; 6×SSC
O	DNA : RNA	≥50	55°C; 4×SSC -or- 42°C; 6×SSC,50% formamide	55°C; 2×SSC
P	DNA : RNA	<50	T <sub>P</sub> *; 6×SSC	T <sub>P</sub> *; 6×SSC
Q	RNA : RNA	≥50	60°C; 4×SSC -or- 45°C; 6×SSC,50% formamide	60°C; 2×SSC
R	RNA : RNA	<50	T <sub>R</sub> *; 4×SSC	T <sub>R</sub> *; 4×SSC

‡ : The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

† : SSPE (1×SSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH7.4) can be substituted for SSC (1×SSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

\*T<sub>B</sub> - T<sub>R</sub> : The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(°C)=2(#of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(°C)=81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41 (%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1×SSC=0.165M).



Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory

5 Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Current Protocols in Molecular Biology, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

10 Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least  
15 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is  
20 determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

25 EXAMPLE

The present invention is embodied in more detail by the following examples, but this embodiment is not intended to restrict the present invention. The basic operations and the enzyme reactions with regard to the DNA recombination are

carried out according to the literature ["Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Laboratory, 1989]. Unless otherwise stated, restrictive enzymes and a variety of modification enzymes to be used were those available from  
5 Takara Shuzo Co., Ltd. The manufacturer's instructions were used for the buffer compositions as well as for the reaction conditions, in each of the enzyme reactions. The cDNA synthesis was carried out according to the literature [Kato, S. et al., Gene 150: 243-250 (1994)].

10 (1) Preparation of Poly(A)<sup>+</sup> RNA

The epidermoid carcinoma cell line KB (ATCC CRL 17), tissues of stomach cancer delivered by the operation, and liver were used for human cells to extract mRNAs. The cell line was cultured by a conventional procedure.

15 After about 1 g of human tissues was homogenized in 20 ml of a 5.5 M guanidinium thiocyanate solution, total mRNAs were prepared in accordance with the literature [Okayama, H. et al., "Methods in Enzymology" Vol. 164, Academic Press, 1987]. These mRNAs were subjected to chromatography using an oligo(dT)-  
20 cellulose column washed with 20 mM Tris-hydrochloric acid buffer solution (pH 7.6), 0.5 M NaCl, and 1 mM EDTA to obtain a poly(A)<sup>+</sup> RNA in accordance with the above-mentioned literature.

(2) Construction of cDNA Library

25 To a solution of 10 µg of the above-mentioned poly(A)<sup>+</sup> RNA in 100 mM Tris-hydrochloric acid buffer solution (pH 8) was added one unit of an RNase-free, bacterium-origin alkaline phosphatase and the resulting solution was allowed to react at 37°C for one hour. After the reaction solution underwent the

phenol extraction followed by the ethanol precipitation, the obtained pellets were dissolved in a mixed solution of 50 mM sodium acetate (pH 6), 1 mM EDTA, 0.1% 2-mercaptoethanol, and 0.01% Triton X-100. Thereto was added one unit of a tobacco-  
5 origin pyrophosphatase (Epicenter Technologies) and the resulting solution at a total volume of 100  $\mu$ l was allowed to react at 37°C for one hour. After the reaction solution underwent the phenol extraction followed by the ethanol precipitation, the thus-obtained pellets were dissolved in  
10 water to obtain a decapped poly(A)<sup>+</sup> RNA solution.

To a solution of the decapped poly(A)<sup>+</sup> RNA and 3 nmol of a DNA-RNA chimeric oligonucleotide (5'-dG-dG-dG-dG-dA-dA-dT-dT-dC-dG-dA-G-G-A-3') in a mixed aqueous solution of 50 mM Tris-hydrochloric acid buffer solution (pH 7.5), 0.5 mM ATP, 5 mM  
15 MgCl<sub>2</sub>, 10 mM 2-mercaptoethanol, and 25% polyethylene glycol were added 50 units of T4 RNA ligase and the resulting solution at a total volume of 30  $\mu$ l was allowed to react at 20°C for 12 hours. After the reaction solution underwent the phenol extraction followed by the ethanol precipitation, the thus-  
20 obtained pellets were dissolved in water to obtain a chimeric oligo-capped poly(A)<sup>+</sup> RNA.

After the vector pKA1 developed by the present inventors (Japanese Patent Kokai Publication No. 1992-117292) was digested with KpnI, an about 60-dT tail was inserted by a  
25 terminal transferase. This product was digested with EcoRV to remove the dT tail at one side and the resulting molecule was used as a vectorial primer.

After 6  $\mu$ g of the previously-prepared chimeric oligo-capped poly(A)<sup>+</sup> RNA was annealed with 1.2  $\mu$ g of the vectorial

primer, the product was dissolved in a mixed solution of 50 mM Tris-hydrochloric acid buffer solution (pH 8.3), 75 mM KCl, 3 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, and 1.25 mM dNTP (dATP + dCTP + dGTP + dTTP), mixed with 200 units of a reverse transferase (GIBCO-BRL), and the resulting solution at a total volume of 20  $\mu$ l was allowed to react at 42°C for one hour. After the reaction solution underwent the phenol extraction followed by the ethanol precipitation, the thus-obtained pellets were dissolved in a mixed solution of 50 mM Tris-hydrochloric acid buffer solution (pH 7.5), 100 mM NaCl, 10 mM MgCl<sub>2</sub>, and 1 mM dithiothreitol. Thereto were added 100 units of EcoRI and the resulting solution at a total volume of 20  $\mu$ l was allowed to react at 37°C for one hour. After the reaction solution underwent the phenol extraction followed by the ethanol precipitation, the obtained pellets were dissolved in a mixed solution of 20 mM Tris-hydrochloric acid buffer solution (pH 7.5), 100 mM KCl, 4 mM MgCl<sub>2</sub>, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 50  $\mu$ g/ml bovine serum albumin. Thereto were added 60 units of *Escherichia coli* DNA ligase and the resulting solution was allowed to react at 16°C for 16 hours. To the reaction solution were added 2  $\mu$ l of 2 mM dNTP, 4 units of *Escherichia coli* DNA polymerase I, and 0.1 unit of *Escherichia coli* DNase H and the resulting solution was allowed to react at 12°C for one hour and then at 22°C for one hour.

Next, the cDNA-synthesis reaction solution was used to transform *Escherichia coli* DH12S (GIBCO-BRL). The transformation was carried out by the electroporation method. A portion of the transformant was inoculated on a 2xYT agar culture medium containing 100  $\mu$ g/ml ampicillin, which was

incubated at 37°C overnight. A colony grown on the culture medium was randomly picked up and inoculated on 2 ml of the 2xYT culture medium containing 100 µg/ml ampicillin, which was incubated at 37°C overnight. The culture medium was centrifuged to separate the cells, from which a plasmid DNA was prepared by the alkaline lysis method. After the plasmid DNA was double-digested with EcoRI and NotI, the product was subjected to 0.8% agarose gel electrophoresis to determine the size of the cDNA insert. In addition, by the use of the obtained plasmid as a template, the sequence reaction using M13 universal primer labeled with a fluorescent dye and Taq polymerase (a kit of Applied Biosystems Inc.) was carried out and the product was analyzed by a fluorescent DNA-sequencer (Applied Biosystems Inc.) to determine the base sequence of the cDNA 5'-terminal of about 400 bp. The sequence data were filed as a homo-protein cDNA bank data base.

### (3) Selection of cDNAs Encoding Proteins Having Transmembrane Domains

The base sequence registered in the homo-protein cDNA bank data base was converted to three frames of amino acid sequences and the presence or absence of an open reading frame (ORF) beginning from the initiation codon. Then, the selection was made for the presence of a signal sequence that is characteristic to a secretory protein at the N-terminal of the portion encoded by ORF. These clones were sequenced from the both 5' and 3' directions by using the deletion method to determine the sequence of the whole base sequence. The hydrophobicity/hydrophilicity profiles were obtained for proteins encoded by ORF by the Kyte-Doolittle method [Kyte, J.

& Doolittle, R. F., J. Mol. Bio. 157: 105-132 (1982)] to examine the presence or absence of a hydrophobic region. In the case in which there is a hydrophobic region of putative transmembrane domain(s) in the amino acid sequence of an encoded protein, this protein was considered as a membrane protein.

(4) Construction of Secretory Signal Detection Vector  
pSSD3

One microgram of pSSD1 carrying the SV40 promoter and a cDNA encoding the protease domain of urokinase [Yokoyama-Kobayashi, M. et al., Gene 163: 193-196 (1995)] was digested with 5 units of BglII and 5 units of EcoRV. Then, after dephosphorylation at the 5' terminal by the CIP treatment, a DNA fragment of about 4.2 kbp was purified by cutting off from the gel of agarose gel electrophoresis.

Two oligo DNA linkers, L1 (5'-GATCCCGGGTCACGTGGGAT-3') and L2 (5'-ATCCACGTGACCCGG-3'), were synthesized and phosphorylated by T4 polynucleotide kinase. After annealing of the both linkers, followed by ligation with the previously-prepared pSSD1 fragment by T4 DNA ligase, *Escherichia coli* JM109 was transformed. A plasmid pSSD3 was prepared from the transformant and the objective recombinant was confirmed by the determination of the base sequence of the linker-inserted fragment. Figure 1 illustrates the structure of the thus-obtained plasmid. The present plasmid vector carries three types of blunt-end formation restriction enzyme sites, SmaI, PmaCI, and EcoRV. Since these cleavage sites are positioned in succession at an interval of 7 bp, selection of an appropriate site in combination of three types of frames for the inserting

sub BP could  
CDNA allows to construct a vector expressing a fusion protein.

#### (5) Functional Verification of Secretory Signal Sequence

Whether the N-terminal hydrophobic region in the secretory protein clone candidate obtained in the above-mentioned steps functions as the secretory signal sequence was verified by the method described in the literature [Yokoyama-Kobayashi, M. et al., Gene 163: 193-196 (1995)]. First, the plasmid containing the target cDNA was cleaved at an appropriate restriction enzyme site that existed at the downstream of the portion expected for encoding the secretory signal sequence. In the case in which this restriction enzyme site was a protruding terminus, the site was blunt-ended by the Klenow treatment or treatment with the mung-bean nuclease. Digestion with HindIII was further carried out and a DNA fragment containing the SV40 promoter and a cDNA encoding the secretory sequence at the downstream of the promoter was separated by agarose gel electrophoresis. This fragment was inserted between the pSSD3 HindIII site and a restriction enzyme site selected so as to match with the urokinase-coding frame, thereby constructing a vector expressing a fusion protein of the secretory signal portion of the target cDNA and the urokinase protease domain.

After *Escherichia coli* (host: JM109) bearing the fusion-protein expression vector was incubated at 37°C for 2 hours in 2 ml of the 2xYT culture medium containing 100 µg/ml ampicillin, the helper phage M13KO7 (50 µl) was added and the incubation was continued at 37°C overnight. A supernatant separated by centrifugation underwent precipitation with polyethylene glycol to obtain single-stranded phage particles. These particles were suspended in 100 µl of 1 mM Tris-0.1 mM

EDTA, pH 8 (TE). Also, there was used as a control a suspension of single-stranded particles prepared in the same manner from the vector pLAI-UPA containing pSSD3 and a full-length cDNA of urokinase [Yokoyama-Kobayashi, M. et al., Gene 163: 193-196 (1995)].

The simian-kidney-origin culture cells, COS7, were incubated at 37°C in the presence of 5% CO<sub>2</sub> in the Dulbecco's modified Eagle's culture medium (DMEM) containing 10% bovine fetus albumin. Into a 6-well plate (Nunc Inc., 3 cm in the well diameter) were inoculated  $1 \times 10^5$  COS7 cells and incubation was carried out at 37°C for 22 hours in the presence of 5% CO<sub>2</sub>. After the culture medium was removed, the cell surface was washed with a phosphate buffer solution and then washed again with DMEM containing 50 mM Tris-hydrochloric acid (pH 7.5) (TDMEM). To the cells were added 1 µl of the single-stranded phage suspension, 0.6 ml of the DMEM culture medium, and 3 µl of TRANSFECTAM<sup>TM</sup> (IBF Inc.) and the resulting mixture was incubated at 37°C for 3 hours in the presence of 5% CO<sub>2</sub>. After the sample solution was removed, the cell surface was washed with TDMEM, 2 ml per well of DMEM containing 10% bovine fetus albumin was added, and the incubation was carried out at 37°C for 2 days in the presence of 5% CO<sub>2</sub>.

To 10 ml of 50 mM phosphate buffer solution (pH 7.4) containing 2% bovine fibrinogen (Miles Inc.), 0.5% agarose, and 1 mM potassium chloride were added 10 units of human thrombin (Mochida Pharmaceutical Co., Ltd.) and the resulting mixture was solidified in a plate of 9 cm in diameter to prepare a fibrin plate. Ten microliters of the culture supernatant of the



transfected COS7 cells were spotted on the fibrin plate, which was incubated at 37°C for 15 hours. The diameter of the thus-obtained clear circle was taken as an index for the urokinase activity. In the case in which a cDNA fragment codes for the amino acid sequence that functions as a secretory signal sequence, a fusion protein is secreted to form a clear circle by its urokinase activity. Therefore, in the case in which a clear circle is not formed, the fusion protein remains as trapped in the membrane and the cDNA fragment is considered to code for a transmembrane domain.

#### (6) Protein Synthesis by In Vitro Translation

The plasmid vector carrying the cDNA of the present invention was utilized for the transcription/translation by the T<sub>N</sub>T rabbit reticulocyte lysate kit (Promega Biotec). In this case, [<sup>35</sup>S]methionine was added and the expression product was labeled with the radioisotope. All reactions were carried out by following the protocols attached to the kit. Two micrograms of the plasmid was allowed to react at 30°C for 90 minutes in total 25 ml of a reaction solution containing 12.5 µl of the T<sub>N</sub>T rabbit reticulocyte lysate, 0.5 µl of the buffer solution (attached to the kit), 2 µl of an amino acid mixture (methionine-free), 2 µl (0.37 MBq/µl) of [<sup>35</sup>S]methionine (Amersham Corporation), 0.5 µl of T7 RNA polymerase, and 20 U of RNasin. To 3 µl of the reaction solution was added 2 µl of an SDS sampling buffer (125 mM Tris-hydrochloric acid suffer solution, pH 6.8, 120 mM 2-mercaptoethanol, 2% SDS solution, 0.025% bromophenol blue, and 20% glycerol) and the resulting solution was heated at 95°C for 3 minutes and then subjected to SDS-polyacrylamide gel electrophoresis. The molecular weight of

the translation product was determined by carrying out the autoradiography.

(7) Expression in COS7

Escherichia coli bearing a vector expressing the protein of the invention was infected with helper phage M13KO7, and single-stranded phage particles were obtained according to the method as stated above. Using the thus obtained phages, each expression vecotr was introduced into simian-kidney-origin culture cells COS7 in the manner as stated above. After incubation at 37 °C for 2 days in the presence of 5 % CO<sub>2</sub>, further incubation was carried out in a medium containing [<sup>35</sup>S]cysteine or [<sup>35</sup>S]methionine for 1 hour. The cells were collected, dissolved and then subjected to SDS-PAGE whereby a band corresponding to the expression product of each protein which is not present in COS7 cells was revealed. In Table 3, the molecular weight of each expression product is shown.

Table 3

20	HP Number	Supernatant of culture	Membrane fraction
		(kDa)	(kDa)
	HP01263	50	-
	HP01299	-	30
	HP01526	-	22
25	HP10230	-	24
	HP10408	-	7
	HP10415	-	45
	HP10424	-	14
	HP10429	-	27
30	HP10432	-	17
	HP10480	-	22

## (8) Clone Examples

&lt;HP01263&gt; (Sequence Number 1, 19, 37)

Determination of the whole base sequence for the cDNA insert of clone HP01263 obtained from the human liver cDNA libraries revealed the structure consisting of a 5'-non-translation region of 36 bp, an ORF of 1149 bp, and a 3'-non-translation region of 316 bp. The ORF codes for a protein consisting of 382 amino acid residues with one transmembrane domain at the N-terminal. Figure 2 depicts the hydrophobicity /hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. The in vitro translation resulted in formation of a translation product of 42 kDa, which is almost consistent with the molecular weight of 42,054 as predicted from the ORF. On expression in COS cells, an expression product of about 50 kDa was observed in the culture supernatant. Therefore, said protein can be understood to be a secreted protein. Application of the rule (-3, -1) as a method for anticipation of a cutting site in a secretion signal sequence suggested that the mature protein would start from methionine at 19 position.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the human  $\alpha$ -2-HS-glycoprotein (SWISS-PROT Accession No. P02765). Table 4 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the human  $\alpha$ -2-HS-glycoprotein (GP). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the

09445558 "120199"

sub  
B2

protein of the present invention. The both proteins possessed  
 a homology of 25.5%. The cysteine position is reserved and this  
 region is analogous to that in cystatins (thiol proteinase  
 inhibitors). There are observed other analogy with histidine-  
 rich glycoprotein (P04196, 30.9%/194 amino acid residues),  
 kininogen (P01045, 24.1%/261 amino acid residues), tyrosine  
 kinase inhibitor (A32827, 24.4%/291 amino acid residues), and  
 so on.

Table 4

10

HP MGLLLPLALCILVLCCGAMSPQALNPSALLSR--GCNDSVLA VAGFALRDINKDRKD  
 .\*\*\* ... \* . . . . \* . \* . . . \* . \* . . .

GP MKSLVLLLCLLAQLWGCHSAPHGPGLIYRQPNCDPETEEAALVAIDYINQNLFW

HP GYVLRRLNRVNDAQEYRRGGLGSLFYLTLDVLETDCHVLRKKAWQDCGMRIFFE-SVYGQC

15

\*\* \*\* \* . . . . . \* . . . . . \* . . . . . \* . . . . .

GP GYKHTLNQIDEVKVWPQQPSGELFEIEIDTLETTCHVLDPTPVARCSVRQLKEHAVEGDC

HP K-AIFYMNNPSRVLYLAAYNCTLRPVSKKKIYMTCPDCPSSIPTDSSNHQVLEAATESLA

. . . . . \* . . . . . \* . . . . . \* . . . . . \* . . . . .

GP DFQLKLKLDGKFSVVY---AKCDSSPDSAEDVRKVCQDCPLLAPLN--DTRVVHAAKAALA

20

HP KYNNTSKQYSLFKVTRASSQWVGPSYFVEYLIKESPC---TKSQASSCSLQSSDSVP

. \* . . . . . \* . . . . . \* . . . . . \* . . . . . \* . . . . .

GP AFNAQNNGSNFQLEEISRAQLV-PLPPSTYVEFTVSGTDCVAKEATEAAKCNLLAEKQY-

HP VGLCKGSLTRTHWEKFVSVTCDFEFESQAPATGSENSAVNQK-PTNLPKVEESQKNTPTPT

\* . \* . . . . \* . . . . . \* . . . . . \* . . . . . \* . . . . .

25

GP -GFCKATLSEKLGGAEVAVTCTVFQTQPVTSQPQPEGANEAVPTPVVDPDAPPSPLGAP

HP DSPSKAGPRGSVQYLPDLDDKNSQEKGPQEAFPVHLDLTNPQGETLDISFLFLEPMEEK

. \* . . . . . \* .

GP GLPPAGSPDSEVLLAAPPQHQLHRAHYDLRHTFMGVVSLGSPSGEVSHPRKTRTVVQPS

HP LVVLPFPKEKARTAECPGPAQNASPLVLPP

30

GP VGAAAGPVVPPCGRIRHFKV

04455-1027-8525-460

Sub  
B2  
Cont'd

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. H57204), but it can not be assessed whether these ESTs with  
5 partial sequences code for the same protein as the protein of the present invention. Hereupon, most of ESTs matching with the present cDNA are available from liver cDNA libraries, whereby the present clone is considered to be expressed specifically in the liver.

10 The present protein, because of being a type-II membrane protein, is considered to exert its function as a receptor on the membrane surface with the C-terminal side exposed outside the cells or after undergoing a processing followed by being excreted in the serum. The present protein, because of bearing  
15 a cystatin-like domain, is considered to possess a proteinase-inhibitor activity as well as many physiological activities in the same manner as for other members of this family. In addition, the present protein, because of being expressed specifically in liver cells, is considered to play a  
20 significant role for maintaining the liver function.

<HP01299> (Sequence Number 2, 20, 38)

Determination of the whole base sequence for the cDNA insert of clone HP01299 obtained from the human liver cDNA libraries revealed the structure consisting of a 5'-non-  
25 translation region of 110 bp, an ORF of 954 bp, and a 3'-non-translation region of 285 bp. The ORF codes for a protein consisting of 317 amino acid residues with two or more transmembrane domains. Figure 3 depicts the hydrophobicity/hydrophilicity profile of the present protein

obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 32 kDa that was almost consistent with the molecular weight of 35,965 predicted from the ORF.

5        The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the rat retinol dehydrogenase (NBRF Accession No. A55884). Table 5 indicates the comparison of the amino acid sequences between the human protein of the present invention  
10    (HP) and the rat retinol dehydrogenase (RN). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and. represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed a homology of 65.3%  
15    among the entire regions.

pub  
B3

09445258.1 201909

[illegible]

The rat retinol dehydrogenase has been found as a  
30 microsomal membrane protein participating in the retinoic acid

biosynthesis in the liver [Chai, X. et al., J. Biol. Chem. 270: 28408-28412 (1995)]. Accordingly, its homologue, the protein of the present invention, is considered to possess a similar function and can be utilized for diagnosis and treatment of diseases caused by the abnormality of this protein.

<HP01347> (Sequence Number 3, 21, 39)

Determination of the whole base sequence for the cDNA insert of clone HP01347 obtained from the human liver cDNA libraries revealed the structure consisting of a 5'-non-translation region of 24 bp, an ORF of 891 bp, and a 3'-non-translation region of 728 bp. The ORF codes for a protein consisting of 296 amino acid residues with one transmembrane domain at the N-terminal. Figure 4 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that the present protein remained in the membrane from the observation that the urokinase secretion was not identified and the urokinase activity was detected on the membrane surface, upon transduction into the COS7 cells of an expression vector in which a HindIII-SacI fragment (treated with the mung-bean nuclease) containing a cDNA fragment encoding the N-terminal 73 amino acid residues in the present protein was inserted at the HindIII-EcoRV site of pSSD3. Therefore, the present protein is considered to be a type-II membrane protein. The in vitro translation resulted in the formation of a translation product of 33 kDa that was almost consistent with the molecular weight of 33,527 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was

sub  
by



analogous to the human HIV envelope glycoprotein gp120-binding C-type lectin (GenBank Accession No. M98457). Table 6 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the human HIV envelope glycoprotein gp120-binding C-type lectin (CL). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed a homology of 85.6% among 284 amino acid residues. There is observed at the downstream of the transmembrane domain a sequence with seven repetition of Ile-Tyr-Gln-Xaa-Leu-Thr-Xaa-Leu-Lys-Ala-Ala-Val-Gly-Glu-Leu-Xaa-Xaa-Xaa-Ser-Lys-Xaa-Gln-Xaa.

Sub  
B4094433-12049  
657027"8524460

Table 6

	HP	MSDSKEPRVQQLGLL-----GCLGHGALVLQLLSFMLLAGVLVAI
		*****.***** *****.***** ****
5	CL	MSDSKEPRLQQLGLLEEEQLRGLGFRQTRGYKSLAGCLGHGPLVLQLLSFTLLAG----L
	HP	LVQVSKVPSSLSQEQSEQDAIYQNLTLKAAVGELSEKSKLQEIQELTLKAAVGELPE
		*****.***** *****
	CL	LVQVSKVPSSISQEQSRQDAIYQNLTLKAAVGELSEKSKLQEIQELTLKAAVGELPE
	HP	KSKLQEIQELTRLKAAVGELPEKSKLQEIQELTRLKAAVGELPEKSKLQEIQELTRL
10		*****.*****.*****
	CL	KSKLQEIQELTRLKAAVGELPEKSKLQEIQELTWLKAAVGELPEKSKMQEIQELTRL
	HP	KAAGVGLPEKSKLQEIQELTELKAAVGELPEKSKLQEIQELTLKAAVGELPDQSKQQ
		***** ***** ***** *****.*****.*****
	CL	KAAGVGLPEKSKQEQEIQELTRLKAAVGELPEKSKQEQEIQELTRLKAAVGELPEKSKQQ
15	HP	QIQELTDLKTAFERLCRHCPKDWTFQGNCFMSNSQRNWHDSVTACQEVRAQLVVIKT
		.*****.*.* *****.***.*****.***.* *****.
	CL	EIQELTLKAAVERLCHPCPWEWTFQGNCFMSNSQRNWHDSITACKEVGAQLVVIKS
	HP	AEEQLPAVLEQWRTQQ
		**** *. *...
20	CL	AEEQNFLQLQSSRSNRFTWMGLSDLNQEGTWQWVDGSPLLPSFKQYWNRGEPNNVGEEDC

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. H90360), but it can not be assessed whether these ESTs with partial sequences code for the same protein as the protein of the present invention.

The present protein, because of being a type-II membrane protein, is considered to exert its function as a receptor on

the membrane surface with the C-terminal side exposed outside the cells or after undergoing a processing followed by being excreted in the serum. Hereupon, the human HIV envelope glycoprotein gp120-binding C-type lectin that is highly homologous with the present protein has been found as a CD4-independent HIV receptor [Curtis, B. M. et al., Proc. Natl. Acad. Sci. USA 89: 8356-8360 (1992)].

<HP01440> (Sequence Number 4, 22, 40)

Determination of the whole base sequence for the cDNA insert of clone HP01440 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 37 bp, an ORF of 594 bp, and a 3'-non-translation region of 98 bp. The ORF codes for a protein consisting of 197 amino acid residues with four transmembrane domains. Figure 5 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 21 kDa that was almost consistent with the molecular weight of 20,822 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the human tumor-associated antigen L6 (SWISS-PROT Accession No. P30408). Table 7 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the human tumor-associated antigen L6 (L6). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed

04455-1009  
051027 32254460

Pub  
B5

B5 could  
a homology of 47.0% among the entire regions.

Table 7

---

5	HP	MCTGKCARCVGLSLITLCLVCIVANALLVPNGETSWTNTNHLSQLQVWLMGGFIGGGLMV
		** .....* .....* .....* .....* .....* .....* .....*
	L6	MCYGKCARCIGHSLVGLALLCIAANILLYFPNGETKYASENHLRFVWFFSGIVGGGLLM
	HP	LCPG---IAAVRAGGKGGCCGAGCCGNRCMLRSVFSSAFGVLGAIYCLSVSGAGLRNGPR
		* *. * . ... **** . **.* **.*... .*. *. **.*... ** .**
10	L6	LLPAFVFIGLEQDDCCGCCGHENCGKRCAMLSSVLAALIGIAGSGYCVIVAALGLAEGPL
	HP	CLMN-GEWGYHFEDTAGAYLLNRTLWDRCEAPPRVVPWNVTLFSLLVAASCLEIVLCGIQ
		** . *.*.* .....* .....* .....* .....* .....* .....* .....*
	L6	CLDSLQWNYTFASTEGQYLLDTSTWSECTEPKHIVEWNVSLFSILLALGGIEFILCLIQ
	HP	LVNATIGVFCGDCRKKQDTPH
15		..*...* .** * ..*.
	L6	VINGVLGGICGFCCSHQQQYDC

---

Furthermore, the search of GenBank using the base sequence  
 20 of the present cDNA revealed that there existed some ESTs  
 possessing the homology of 90% or more and also containing the  
 initiation codon (for example, Accession No. T55097), but many  
 sequences were not distinct and the same ORF as that in the  
 present cDNA was not identified.

25 The human tumor-associated antigen L6 is a member of a  
 membrane antigen TM4 superfamily proteins which are expressed  
 in large quantities on the surface of human tumor cells  
 [Marken, J. S. et al., Proc. Natl. Acad. Sci. USA 89: 3503-3507  
 (1992)]. Since these membrane antigens are expressed  
 30 specifically on some specified cells or cancer cells,

antibodies against these antigens, if constructed, are useful for a variety of diagnoses and as carriers for the drug delivery. In addition, the cells in which genes of these membrane antigens are transduced and the membrane antigens are expressed are applicable for detection of the corresponding ligands and so on.

<HP01526> (Sequence Number 5, 23, 41)

Determination of the whole base sequence for the cDNA insert of clone HP01526 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 83 bp, an ORF of 666 bp, and a 3'-non-translation region of 573 bp. The ORF codes for a protein consisting of 221 amino acid residues with a hydrophobic region of putative six transmembrane domains. Figure 6 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 23 kDa that was almost consistent with the molecular weight of 25,030 predicted from the ORF.

20 The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the mouse interstitial cell protein (GenBank Accession No. X96618). Table 8 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the mouse interstitial cell protein (MM). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed

a homology of 79.6% among the entire regions.

Table 8

---

5	HP	MEAGGF LDSLIYGACVVFTLGMFSAGLSDLRHMTRSVDNVQFLPFLTTEVNNLGWLSY
		***** **.. .***.*****.*****.*****.*****.*****.*****.*****
	MM	MEAGGVADSF LSSACVLFTLGMFSTGLSDLREMQRTRSVDN IQFLPFLT TDVNNLSWLSY
	HP	GALKGDGILIVVNTVGAALQ TLYILAYLHYCP RKR VLLQTATLLGVLLLG YGYFWLLVP
		*.*****.**.**.***.*****.*****.*.*.*****.*****.*****
10	MM	GVLKGDGTLIIIVNSVGA VLQ TLYILAYLHYSPQKHGVLLQTATLLAVLLLG YGYFWLLVP
	HP	NPEARLQQLGLFCSVFTISMYLSPLADLAKVIQTKSTQCLSYPLTIATLLTSASWC LYGF
		.*****.*****.*****.*****.*****.*****.*****.*****.*****
	MM	DLEARLQQLGLFCSVFTISMYLSPLADLAKIVQTKSTQRLSFSLTIATLFCSASWSIYGF
	HP	RLRDPYIMVSNFPGIVTSFIRFWLFWKYPQEQRNYWLLQT
15		***** *.*.***.***.***. ** ***.***.*.*****
	MM	RLRDPYIAVPNLP GILTS LIRLGLFCKYPPEQDRKYRL LQT

---

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more and also containing the initiation codon (for example, Accession No. H02682), but many sequences were not distinct and the same ORF as that in the present cDNA was not identified.

The mouse interstitial cell protein has been cloned as a membrane protein that is expressed with highly increasing in interstitial cells stimulated by a cytokine [Tagoh, H. et al., Biochem. Biophys. Res. Commun. 221: 744-749 (1996)]. Since these membrane proteins are expressed specifically on some specified cells and cancer cells, antibodies against these

044553-12049  
657027-332446

proteins, if constructed, are useful for a variety of diagnoses and as carriers for the drug delivery. In addition, the cells in which genes of these membrane antigens are transduced and the membrane antigens are expressed are applicable for  
5 detection of the corresponding ligands and so on.

<HP10230> (Sequence Number 6, 24, 42)

Determination of the whole base sequence for the cDNA insert of clone HP10230 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-  
10 translation region of 190 bp, an ORF of 756 bp, and a 3'-non-translation region of 2099 bp. The ORF codes for a protein consisting of 251 amino acid residues with at least one transmembrane domain. Figure 7 depicts the hydrophobicity/hydrophilicity profile of the present protein  
15 obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 30 kDa that was almost consistent with the molecular weight of 28,800 predicted from the ORF.

The search of the protein data base using the amino acid  
20 sequence of the present protein revealed that the protein was analogous to the nematode hypothetical protein F25D7.1 (GenBank Accession No. Z78418). Table 9 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the nematode hypothetical protein F25D7.1  
25 (CE). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed a homology of 49.8% among the entire regions.

044523 1209  
06F02T 03254460

pub  
B7

Table 9

---

HS MSDIGDWFRSIPAITRYWFAATVAVPLVVGKLGGLISPAYLFL-WPEAFLYRFQIWRPITAT  
       \*..... \*\* .\*\*\*\*\* \*.. .\*\*\*.\*\*\*. ....\*\* \* . . .\*\*\*\*\*.\*\*  
 5 CE MDLENFLLGIPIVTRYWFLASTIIPLLGRFGFINVQWMFLQW-DLVVNKFQFWRPLTAL  
 HS FYFPVGP GTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNW-ICIVITGLAMDM  
       .\*\*\*.\* \*\*\* \*. \*\*\*\*\*.\*\*\*. ....\*\* .\*\*\*\*\*.\*\*\* \* . .\*\*\*.  
 CE IYYPVTPQTGFHWLMMCYFLYNYSKALESETYRGRSADYLFMLIFNWFFCSGLC-MALDI  
 HS QLLMIPLIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNL  
 10       \*. \*...\*\*\*\*\* \*.\*. \* \*\*\*\*\* \*\* \* \*\*\*\*\*. \*\*\* .. \*. .\*\*\*.\* \*  
 CE YFLLPEPMVISVLYVWCQVNKDTIVSFWFGMRFPARYLPWVLWGFNAVLRRGGGTNELVGIL  
 HS VGHLYFFLMFRYPMDLGGRNFLSTPQFLYRWLPSSRRGGVSGFGVPPASMRRAADQNGGGG  
       \*\*\* \*\*\*. ....\* \* ...\*\*\*.\*\*\*. \*. \*\*.\* \* \* \* \*  
 CE VGHAYFFVALKYPDEYGV-DLISTPEFLHRLIPDEDGGIHG---QDGNIRGARQQPRG--  
 15 HS RHNW--GQGFRLGDQ  
       \* \* \* \* \*  
 CE -HQWPGGVGARLGCN

---

20 Furthermore, the search of GenBank using the base sequence  
 of the present cDNA revealed that there existed some ESTs  
 possessing the homology of 90% or more and also containing the  
 initiation codon (for example, Accession No. W01493), but many  
 sequences were not distinct and the same ORF as that in the  
 25 present cDNA was not identified.

<HP10389> (Sequence Number 7, 25, 43)

Determination of the whole base sequence for the cDNA  
 insert of clone HP10389 obtained from the human epidermoid  
 carcinoma cell line KBc cDNA libraries revealed the structure  
 30 consisting of a 5'-non-translation region of 62 bp, an ORF of

044553-12046



321 bp, and a 3'-non-translation region of 270 bp. The ORF codes for a protein consisting of 106 amino acid residues with a hydrophobic region of putative two transmembrane domains. Figure 8 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 12 kDa that was almost consistent with the molecular weight of 11,528 predicted from the ORF.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was not analogous to any of known proteins. Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. H70816), but many sequences were not distinct and the same ORF as that in the present cDNA was not identified.

<HP10408> (Sequence Number 8, 26, 44)

Determination of the whole base sequence for the cDNA insert of clone HP10408 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 74 bp, an ORF of 237 bp, and a 3'-non-translation region of 128 bp. The ORF codes for a protein consisting of 78 amino acid residues with a putative signal sequence at the N-terminal as well as a sequence of one putative interior transmembrane domain. Figure 9 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that the present protein remained in the membrane from the observation that the urokinase secretion was not identified

upon transduction into the COS7 cells of an expression vector in which a HindIII-BglII fragment (after the Klenow treatment) containing a cDNA fragment encoding the N-terminal 70 amino acid residues in the present protein was inserted at the  
5 HindIII-EcoRV site of pSSD3. The in vitro translation resulted in the formation of a translation product of 9 kDa that was almost consistent with the molecular weight of 8,396 predicted from the ORF.

Furthermore, the search of GenBank using the base sequence  
10 of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. T94049), but they were shorter than the present cDNA and any molecule containing the initiation codon was not identified.

15 <HP10412> (Sequence Number 9, 27, 45)

Determination of the whole base sequence for the cDNA insert of clone HP10412 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 55 bp, an ORF of 945 bp, and a 3'-non-translation region of 131 bp. The ORF codes for a protein  
20 consisting of 314 amino acid residues with one transmembrane domain at the N-terminal. Figure 10 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that  
25 the present protein remained in the membrane from the observation that the urokinase secretion was not identified upon transduction into the COS7 cells of an expression vector in which a HindIII-ApaI fragment (treated with mung-bean nuclease) containing a cDNA fragment encoding the N-terminal 65

amino acid residues in the present protein was inserted at the HindIII-EcoRV site of pSSD3. The in vitro translation resulted in the formation of a translation product of 44 kDa that was somewhat larger than the molecular weight of 35,610 predicted  
5 from the ORF.

sub B8  
667027 8529460  
The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the nematode hypothetical protein of 28.5 kDa (SWISS-PROT Accession No. P34623). Table 10 indicates the  
10 comparison of the amino acid sequences between the human protein of the present invention (HP) and the nematode hypothetical protein of 28.5 kDa (CE). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino  
15 acid residue analogous to that in the protein of the present invention. The both proteins possessed a homology of 42.8% in the C-terminal region of 243 amino acid residues.

Table 10

# THE UNIVERSITY OF CHICAGO

25 No. T09311), but it can not be assessed whether these ESTs with

<HP10413> (Sequence Number 10, 28, 46)

30 insert of clone HP10413 obtained from the human stomach cancer

cdna libraries revealed the structure consisting of a 5'-non-translation region of 78 bp, an ORF of 588 bp, and a 3'-non-translation region of 1209 bp. The ORF codes for a protein consisting of 195 amino acid residues with one transmembrane domain at the N-terminal. Figure 11 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that the present protein remained in the membrane from the observation that the urokinase secretion was not identified upon transduction into the COS7 cells of an expression vector in which a HindIII-PmaCI fragment containing a cDNA fragment encoding the N-terminal 65 amino acid residues in the present protein was inserted at the HindIII-PmaCI site of pSSD3. The in vitro translation resulted in the formation of a translation product of 28 kDa that was somewhat larger than the molecular weight of 21,671 predicted from the ORF.

The search of the protein/data base using the amino acid sequence of the present protein revealed that the protein was analogous to the swine steroidal membrane-binding protein (GenBank Accession No. X99714). Table 11 indicates the comparison of the amino acid sequences between the human protein of the present invention (HP) and the swine steroidal membrane-binding protein (SS). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The both proteins possessed a homology of 96.4% among the entire regions.

044558 "120199" 85254460

Sub  
B9

Table 11

---

	HP	MAAEDVVATGADPSDLESGGLLHEIFTSPNLNLLGLCIFLLYKIVRGDQPAASGDSDDD
		*****.*****.**.***** *****
5	SS	MAAEDVAATGADPSELEGGGLLHEIFTSPNLNLLGLCIFLLYKIVRGDQPAAS-DSDDD
	HP	EPPPLPRLKRRDFTPaelRRFDGVQDPRILMAINGKVFDVTKGRKFYGPegPYGVFAGRD
		*****
	SS	EPPPLPRLKRRDFTPaelRRFDGVQDPRILMAINGKVFDVTKGRKFYGPegPYGVFAGRD
	HP	ASRGLATFCLDKEALKDEYDDLSDLTAAQETLSDWESQFTFKYHHVGKLLKEGEEPTVY
10		*****.*****.**.*****
	SS	ASRGLATFCLDKEALKDEYDDLSDLTPAQETLNDWDSQFTFKYHHVGKLLKEGEEPTVY
	HP	SDEEEPKDESARKND
		*****
	SS	SDEEEPKDESARKND

---

15

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. AA021062), but many sequences were not distinct and the same ORF as that in the present cDNA was not identified.

<HP10415> (Sequence Number 11, 29, 47)

Determination of the whole base sequence for the cDNA insert of clone HP10415 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 71 bp, an ORF of 1389 bp, and a 3'-non-translation region of 103 bp. The ORF codes for a protein consisting of 462 amino acid residues with one transmembrane domain at the N-terminal. Figure 12 depicts the hydrophobicity/hydrophilicity profile of the present protein

044553 "3534460

obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 48 kDa that was somewhat smaller than the molecular weight of 52,458 predicted from the ORF.

5        The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the cytochrome P450 as exemplified by the simian cytochrome P450IIIA8 (SWISS-PROT Accession No. P33268). Table 12 indicates the comparison of the amino acid sequences between  
10 the human protein of the present invention (HP) and the simian cytochrome P450IIIA8 (CP). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present invention. The  
15 both proteins possessed a homology of 21.3% among the entire regions.

sub  
B10

0944558-100499

Table 12

	HP	MLDFAIFAVTFLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVN-SGSLHEF
		.....*.....*.....*
5	CP	MDLIPDLAVETWLLAVTLVLLYLYGTHSHGLFKKLGIPGPTPLPLGNILSYRKGFWTF
	HP	LVLNHERYGPVVSFWFGRRLVVS LGTVDLKQHINPNKTLDPFETMLK-SLLRYQSGGGS
		.....** * ..* ..* .....* ..*.....*
	CP	DMECYKKYKGVWGFYDGRQPVLAITDPNMIK-TVLVKECYSVFTNRRPFGPVGFMKNAIS
	HP	VSEN----HMRKKLYENGVTDSLKSNFALLKLSEELLDKWSYPET-QHVPLSQHMLGF
10		..* ..* * .....** ..* ..* ..* ..* ..* ..* ..* ..* ..*
	CP	IAEDEEWKRIRSLSPFTTSGKLKEMVPIIAKYGDVLVRNLRREAETGKPVTLKDVFGAY
	HP	AMKSVTQMVMG-----STF-EDDQEVIRFQKNHGTWVSEIGKGFLDGSLD--KNM
		..* ..* ..* .....* .....* ..* ..* ..* ..*
	CP	SMDVITSTSFGVNIDSLNNPQDPFVENTKKLLRFDFLDPFFLSITIFPFIPILEVNLIS
15	HP	TRKKQYEDALMQ-LESVLRNIIKE-RKGR-NFSQHIF-----IDSLVQGNLNDQQILED
		.. ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..*
	CP	IFPREVTSFLRKS VKRIKESRLKDTQKHRVDFLQLMIDSQNSKETESHKALSDLELVAQS
	HP	MIFSLASCIITAKLCTWAICFLTSEEVQKKLYEEINQVF-GNGPVTPEKIEQLRYCQHV
		..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..*
20	CP	IIFIFAGYETTSSVLSFIIYELATHPDVQKQLQEEIDTVLPNKAPPTYDTVLQMEYLDMV
	HP	LCETVRTAKLTPVSAQLQDIEGKIDRFIIPRETLLVLYALGVVLQDPNTWPSPHKFDPDFR
		..**.* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..*
	HP	VNETLRIFPIAMRLERVCKKDVEINGIFIPKGVVVMIPSYALHHDPKYWPEPEKFLPERF
	HP	----DDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLSVEGQVIETKYE
25		..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..* ..*
	CP	SKKNNDNIDPYIYTPFG-SGPRNCIGMRFALNMKLAIIRVLQNFSEFKPCKETQIPLKLR
	HP	LVTSSREEAWITVSKRY
		*
	CP	LGLLQTEKPIVLKIESRDGTVSGA

30

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs



possessing the homology of 90% or more (for example, Accession No. AA381169), but it can not be assessed whether these ESTs with partial sequences code for the same protein as the protein of the present invention.

- 5       The cytochrome P450 participates in the drug metabolism and can be utilized as a catalyst in organic synthesis reactions such as oxidation and so on.

<HP10419> (Sequence Number 12, 30, 48)

- 10       Determination of the whole base sequence for the cDNA insert of clone HP10419 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 170 bp, an ORF of 744 bp, and a 3'-non-translation region of 1116 bp. The ORF codes for a protein consisting of 247 amino acid residues with a hydrophobic region  
15       of putative seven transmembrane domains. Figure 13 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method.

- 20       The search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. AA340663), but it can not be assessed whether these ESTs with partial sequences code for the same protein as the protein of the present invention.

<HP10424> (Sequence Number 13, 31, 49)

- 25       Determination of the whole base sequence for the cDNA insert of clone HP10424 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 97 bp, an ORF of 342 bp, and a 3'-non-translation region of 54 bp. The ORF codes for a protein

consisting of 113 amino acid residues with one transmembrane domain at the N-terminal. Figure 14 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that the present protein remained in the membrane from the observation that the urokinase secretion was not identified upon transduction into the COS7 cells of an expression vector in which a HindIII-AccI fragment (after the Klenow treatment) containing a cDNA fragment encoding the N-terminal 58 amino acid residues in the present protein was inserted at the HindIII-SmaI site of pSSD3. The in vitro translation resulted in the formation of a translation product of 14 kDa that was somewhat larger than the molecular weight of 12,784 predicted from the ORF.

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. AA401979), but it can not be assessed whether these ESTs with partial sequences code for the same protein as the protein of the present invention.

<HP10428> (Sequence Number 14, 32, 50)

Determination of the whole base sequence for the cDNA insert of clone HP10428 obtained from the human epidermoid carcinoma cell line KBc cDNA libraries revealed the structure consisting of a 5'-non-translation region of 287 bp, an ORF of 1098 bp, and a 3'-non-translation region of 659 bp. The ORF codes for a protein consisting of 365 amino acid residues with a hydrophobic region of putative nine transmembrane domains. Figure 15 depicts the hydrophobicity/hydrophilicity profile of

the present protein obtained by the Kyte-Doolittle method. The result of the in vitro translation did not reveal the formation of distinct bands and only revealed the formation of smeary bands at the high-molecular-weight position.

5        The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was analogous to the baker's yeast hypothetical membrane protein YML038c (NBRF Accession No. S49741). Table 13 indicates the comparison of the amino acid sequences between the human  
10 protein of the present invention (HP) and the baker's yeast hypothetical membrane protein YML038c (SC). - represents a gap, \* represents an amino acid residue identical to that in the protein of the present invention, and . represents an amino acid residue analogous to that in the protein of the present  
15 invention. The both proteins possessed a homology of 26.3% among the N-terminal region of 281 amino acid residues.

Sub  
B11

0944553-120160

Table 13

---

	HP	MGRWALDVAFLWKAVLTLGLVL-LYYCFSIGITFYNKWL-----TKSPHFPLFMTMLHLA
		*... *... *... *... *... *... *... *... *
5	SC	MNRTVFLAFVFGWYFCS-IALSIYNRWMFDPKDLGIGYPVLVTTFHQA
	HP	VIFLPSALSRAIVQ---CSSHRARVVLVSWADYLRRVAPTALATALDVGLSNWSFLYVTVS
		...*... *... *... *... *... *... *... *... *
	SC	TLWLLSGIYIKLRHKPVKNVLRKNNGFNWSFLLKFLPTAVASAGDIGLSNVSFQYVPLT
	HP	LYTMTKSSAVLFILIFSLIFKLEEL--RAALVLVLLIAGGLFMF-----TYKSTQ-FN
10		...*... *... *... *... *... *... *... *... *
	SC	IYTIKSSSIAFVLLFGCIFKLEKFWKLALSVIIMFVGVALMVFKPSDSTSTKNDQALV
	HP	VEGFALVLGASFIGGIRWTLTQMLLQKAELGLQNPIDTMFHLQPLMFLGLFPLFAVFEGL
		*... *... *... *... *... *... *... *... *
	SC	IFGSFLVLASSCLSGLRWVYTQMLLRNNPIQTNTAAVEES-DGALFTENEDNVDNEPVV
15	HP	HLSTSEKIFRFQDT-GLLLRVLGSFLGILAFGLGFSEFLVSRSSLTLSIAGIFKEV
		*... *... *... *... *... *... *... *... *
	SC	NLANNKMLENFGESKPHPIHTIHQ--LAPIMGITLLLS-LLVEKPPFGIFS-SSIFRLD
	HP	CTLLLAHLLGDQISLLNLWLGFAICLSGISLHVALKALHSRGDGGPKALKGLGSSPDLEL
20	SC	TSNGGVGTETTIVLSIVRGIVLLILPGFAVFLLTICEFSILEQTPVLTVSIVGIVKELLTV
	HP	LLRSSQREEGDNEEEYFVAQGQQ
	SC	IFGIIILSERLSGFYNWLGMIIIMADVCIYNYFRYKQDLLQKYHSVSTQDNRNELKGFQD

---

Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. AA018345), but it can not be assessed whether these ESTs

with partial sequences code for the same protein as the protein of the present invention.

<HP10429> (Sequence Number 15, 33, 51)

Determination of the whole base sequence for the cDNA  
5 insert of clone HP10429 obtained from the human stomach cancer  
cDNA libraries revealed the structure consisting of a 5'-non-  
translation region of 156 bp, an ORF of 681 bp, and a 3'-non-  
translation region of 206 bp. The ORF codes for a protein  
consisting of 226 amino acid residues with four transmembrane  
10 domains. Figure 16 depicts the hydrophobicity/hydrophilicity  
profile of the present protein obtained by the Kyte-Doolittle  
method. The in vitro translation resulted in the formation of  
a translation product of 25 kDa that was almost consistent with  
the molecular weight of 25,321 predicted from the ORF.

15 The search of the protein data base using the amino acid  
sequence of the present protein revealed that the protein was  
not analogous to any known proteins. Furthermore, the search of  
GenBank using the base sequence of the present cDNA revealed  
that there existed some ESTs possessing the homology of 90% or  
20 more (for example, Accession No. AA315933), but it can not be  
assessed whether these ESTs with partial sequences code for the  
same protein as the protein of the present invention.

<HP10432> (Sequence Number 16, 34, 52)

Determination of the whole base sequence for the cDNA  
25 insert of clone HP10429 obtained from the human liver cDNA  
libraries revealed the structure consisting of a 5'-non-  
translation region of 28 bp, an ORF of 390 bp, and a 3'-non-  
translation region of 554 bp. The ORF codes for a protein  
consisting of 129 amino acid residues with a signal-like

sequence at the N-terminal and one interior transmembrane domain. Therefore, the present protein is considered to be a type-I membrane protein. Figure 17 depicts the hydrophobicity/hydrophilicity profile of the present protein  
5 obtained by the Kyte-Doolittle method.

The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was not analogous to any known proteins. Furthermore, the search of GenBank using the base sequence of the present cDNA revealed  
10 that there existed some ESTs possessing the homology of 90% or more (for example, Accession No. T74424), but the same ORF as that in the present cDNA was not identified.

<HP10433> (Sequence Number 17, 35, 53)

Determination of the whole base sequence for the cDNA  
15 insert of clone HP10433 obtained from the human liver cDNA libraries revealed the structure consisting of a 5'-non-translation region of 72 bp, an ORF of 492 bp, and a 3'-non-translation region of 131 bp. The ORF codes for a protein  
20 domain at the N-terminal. Figure 18 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. It was indicated that the present protein remained in the membrane from the observation that the urokinase secretion was not identified  
25 upon transduction into the COS7 cells of an expression vector in which a HindIII-Eco81I fragment (treated with the mung-bean nuclease) containing a cDNA fragment encoding the N-terminal 137 amino acid residues in the present protein was inserted at the HindIII-EcoRV site of pSSD3. Therefore, the present protein

is considered to be a type-II membrane protein. The in vitro translation resulted in the formation of a translation product of 21 kDa that was almost consistent with the molecular weight of 18,617 predicted from the ORF.

5       The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was not analogous to any known proteins. Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or  
10 more (for example, Accession No. H84693), but many sequences are not distinct and the same ORF as that in the present cDNA was not identified.

<HP10480> (Sequence Number 18, 36, 54)

Determination of the whole base sequence for the cDNA  
15 insert of clone HP10480 obtained from the human stomach cancer cDNA libraries revealed the structure consisting of a 5'-non-translation region of 79 bp, an ORF of 582 bp, and a 3'-non-translation region of 1253 bp. The ORF codes for a protein consisting of 193 amino acid residues with four transmembrane  
20 domains. Figure 19 depicts the hydrophobicity/hydrophilicity profile of the present protein obtained by the Kyte-Doolittle method. The in vitro translation resulted in the formation of a translation product of 23 kDa that was somewhat larger than the molecular weight of 21,445 predicted from the ORF.

25       The search of the protein data base using the amino acid sequence of the present protein revealed that the protein was not analogous to any known proteins. Furthermore, the search of GenBank using the base sequence of the present cDNA revealed that there existed some ESTs possessing the homology of 90% or

more (for example, Accession No. W93606), but many sequences are not distinct and the same ORF as that in the present cDNA was not identified.

The present invention provides human proteins having  
5 transmembrane domains and cDNAs encoding said proteins. All of the proteins of the present invention are putative proteins controlling the proliferation and differentiation of the cells, because said proteins exist on the cell membrane. Therefore, the proteins of the present invention can be used as  
10 pharmaceuticals or as antigens for preparing antibodies against said proteins. Furthermore, said DNAs can be used for the expression of large amounts of said proteins. The cells expressing large amounts of membrane proteins with transfection of these membrane protein genes can be applied to the detection  
15 of the corresponding ligands, the screening of novel low-molecular medicines, and so on.

In addition to the activities and uses described above, the polynucleotides and proteins of the present invention may exhibit one or more of the uses or biological activities  
20 (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies  
25 or vectors suitable for introduction of DNA).

#### Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for



analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as

5 molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA

10 sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for

15 examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in

20 a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

25 The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in

assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of  
5 tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which  
10 binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of  
15 being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A  
20 Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

#### 25 Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source

and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

### Cytokine and Cell Proliferation/Differentiation

#### Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H.

Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 5 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of 10 spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse 15 and human Interferon  $\gamma$ , Schreiber, R.D. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without 20 limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 25 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 -Nordan, R. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et

al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined

immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be  
5 caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis  
10 viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

15 Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis,  
20 insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or  
25 other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be

possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2

activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function in vivo on the development of that disease.



Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease.

The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B

lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected ex vivo with an expression

vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection in vivo.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I  $\alpha$  chain protein and  $\beta_2$  microglobulin protein or an MHC class II  $\alpha$  chain protein and an MHC class II  $\beta$  chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a

T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- 5        Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays
- 10    for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol.
- 15    140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowman et al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988;
- 20    Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

- Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that
- 25    affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J.J. and Brunswick, M. In Current Protocols in Immunology. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John

Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in:

Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

#### Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently

of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney,

- M.G. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In Culture of Hematopoietic Cells. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

#### Tissue Growth Activity

- 20 A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.
- 25 A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the



invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue

formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders,

such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

5 Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

10 It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular  
15 endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

20 A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for  
25 promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin  $\alpha$  family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- $\beta$  group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of

the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

5       The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale  
10 et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or  
15 chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell  
20 population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of  
25 infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell

population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

#### Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A

protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses).

Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors

of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

#### Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis,



complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be  
5 useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

#### Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of  
10 the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary  
15 to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth

#### 20 Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi  
25 and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in

bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

09445258 "120199

## CLAIMS

1. A protein comprising an amino acid sequence selected from the group consisting of the amino acid sequences of SEQ  
5 ID NOS: 1 to 18.

2. A DNA encoding the protein according to claim 1.

3. A cDNA comprising a nucleotide sequence selected from the group consisting of the nucleotide sequences of SEQ  
10 ID NOS: 19 to 36.

4. A cDNA according to claim 3, which comprises a nucleotide sequence selected from the group consisting of the  
15 nucleotide sequences of SEQ ID NOS: 37 to 54.

5. An expression vector capable of in vitro translating the DNA according to any of claims 2 to 4 or expressing said DNA in an eukaryotic cell.  
20

6. A transformed eukaryotic cell capable of expressing the DNA according to any of claims 2 to 4 to produce the protein according to claim 1.

044553-460

add a

add  
B16

## Sequence Table

## (2) INFORMATION FOR SEQ ID NO: 1:

## 5 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 382

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

10 (iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Liver

15 (D) CLONE NAME: HP01263

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

20 Met Gly Leu Leu Leu Pro Leu Ala Leu Cys Ile Leu Val Leu Cys Cys  
 1 5 10 15  
 Gly Ala Met Ser Pro Pro Gln Leu Ala Leu Asn Pro Ser Ala Leu Leu  
 20 25 30  
 Ser Arg Gly Cys Asn Asp Ser Asp Val Leu Ala Val Ala Gly Phe Ala  
 35 40 45  
 25 Leu Arg Asp Ile Asn Lys Asp Arg Lys Asp Gly Tyr Val Leu Arg Leu  
 50 55 60  
 Asn Arg Val Asn Asp Ala Gln Glu Tyr Arg Arg Gly Gly Leu Gly Ser  
 65 70 75 80  
 Leu Phe Tyr Leu Thr Leu Asp Val Leu Glu Thr Asp Cys His Val Leu  
 30 85 90 95  
 Arg Lys Lys Ala Trp Gln Asp Cys Gly Met Arg Ile Phe Phe Glu Ser  
 100 105 110  
 Val Tyr Gly Gln Cys Lys Ala Ile Phe Tyr Met Asn Asn Pro Ser Arg  
 115 120 125  
 35 Val Leu Tyr Leu Ala Ala Tyr Asn Cys Thr Leu Arg Pro Val Ser Lys  
 130 135 140  
 Lys Lys Ile Tyr Met Thr Cys Pro Asp Cys Pro Ser Ser Ile Pro Thr  
 145 150 155 160

09445253 120199

Sub  
 a2

84

Asp Ser Ser Asn His Gln Val Leu Glu Ala Ala Thr Glu Ser Leu Ala  
 165 170 175  
 Lys Tyr Asn Asn Glu Asn Thr Ser Lys Gln Tyr Ser Leu Phe Lys Val  
 180 185 190  
 5 Thr Arg Ala Ser Ser Gln Trp Val Val Gly Pro Ser Tyr Phe Val Glu  
 195 200 205  
 Tyr Leu Ile Lys Glu Ser Pro Cys Thr Lys Ser Gln Ala Ser Ser Cys  
 210 215 220  
 Ser Leu Gln Ser Ser Asp Ser Val Pro Val Gly Leu Cys Lys Gly Ser  
 10 225 230 235 240  
 Leu Thr Arg Thr His Trp Glu Lys Phe Val Ser Val Thr Cys Asp Phe  
 245 250 255  
 Phe Glu Ser Gln Ala Pro Ala Thr Gly Ser Glu Asn Ser Ala Val Asn  
 260 265 270  
 15 Gln Lys Pro Thr Asn Leu Pro Lys Val Glu Glu Ser Gln Gln Lys Asn  
 275 280 285  
 Thr Pro Pro Thr Asp Ser Pro Ser Lys Ala Gly Pro Arg Gly Ser Val  
 290 295 300  
 Gln Tyr Leu Pro Asp Leu Asp Asp Lys Asn Ser Gln Glu Lys Gly Pro  
 20 305 310 315 320  
 Gln Glu Ala Phe Pro Val His Leu Asp Leu Thr Thr Asn Pro Gln Gly  
 325 330 335  
 Glu Thr Leu Asp Ile Ser Phe Leu Phe Leu Glu Pro Met Glu Glu Lys  
 340 345 350  
 25 Leu Val Val Leu Pro Phe Pro Lys Glu Lys Ala Arg Thr Ala Glu Cys  
 355 360 365  
 Pro Gly Pro Ala Gln Asn Ala Ser Pro Leu Val Leu Pro Pro  
 370 375 380

30

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 317

(B) TYPE: Amino acid

35

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Liver  
 (D) CLONE NAME: HP01299

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Trp Leu Tyr Leu Ala Ala Phe Val Gly Leu Tyr Tyr Leu Leu His  
 1 5 10 15  
 10 Trp Tyr Arg Glu Arg Gln Val Val Ser His Leu Gln Asp Lys Tyr Val  
 20 25 30  
 Phe Ile Thr Gly Cys Asp Ser Gly Phe Gly Asn Leu Leu Ala Arg Gln  
 35 40 45  
 Leu Asp Ala Arg Gly Leu Arg Val Leu Ala Ala Cys Leu Thr Glu Lys  
 15 50 55 60  
 Gly Ala Glu Gln Leu Arg Gly Gln Thr Ser Asp Arg Leu Glu Thr Val  
 65 70 75 80  
 Thr Leu Asp Val Thr Lys Met Glu Ser Ile Ala Ala Ala Thr Gln Trp  
 85 90 95  
 20 Val Lys Glu His Val Gly Asp Arg Gly Leu Trp Gly Leu Val Asn Asn  
 100 105 110  
 Ala Gly Ile Leu Thr Pro Ile Thr Leu Cys Glu Trp Leu Asn Thr Glu  
 115 120 125  
 Asp Ser Met Asn Met Leu Lys Val Asn Leu Ile Gly Val Ile Gln Val  
 25 130 135 140  
 Thr Leu Ser Met Leu Pro Leu Val Arg Arg Ala Arg Gly Arg Ile Val  
 145 150 155 160  
 Asn Val Ser Ser Ile Leu Gly Arg Val Ala Phe Phe Val Gly Gly Tyr  
 165 170 175  
 30 Cys Val Ser Lys Tyr Gly Val Glu Ala Phe Ser Asp Ile Leu Arg Arg  
 180 185 190  
 Glu Ile Gln His Phe Gly Val Lys Ile Ser Ile Val Glu Pro Gly Tyr  
 195 200 205  
 Phe Arg Thr Gly Met Thr Asn Met Thr Gln Ser Leu Glu Arg Met Lys  
 35 210 215 220  
 Gln Ser Trp Lys Glu Ala Pro Lys His Ile Lys Glu Thr Tyr Gly Gln  
 225 230 235 240  
 Gln Tyr Phe Asp Ala Leu Tyr Asn Ile Met Lys Glu Gly Leu Leu Asn

0944358 "120199"

86

245 250 255  
 Cys Ser Thr Asn Leu Asn Leu Val Thr Asp Cys Met Glu His Ala Leu  
 260 265 270  
 Thr Ser Val His Pro Arg Thr Arg Tyr Ser Ala Gly Trp Asp Ala Lys  
 5 275 280 285  
 Phe Phe Phe Ile Pro Leu Ser Tyr Leu Pro Thr Ser Leu Ala Asp Tyr  
 290 295 300  
 Ile Leu Thr Arg Ser Trp Pro Lys Pro Ala Gln Ala Val  
 305 310 315

10

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 296

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Liver

(D) CLONE NAME: HP01347

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Ser Asp Ser Lys Glu Pro Arg Val Gln Gln Leu Gly Leu Leu Gly  
 1 5 10 15  
 Cys Leu Gly His Gly Ala Leu Val Leu Gln Leu Leu Ser Phe Met Leu  
 30 20 25 30  
 Leu Ala Gly Val Leu Val Ala Ile Leu Val Gln Val Ser Lys Val Pro  
 35 40 45  
 Ser Ser Leu Ser Gln Glu Gln Ser Glu Gln Asp Ala Ile Tyr Gln Asn  
 50 55 60  
 35 Leu Thr Gln Leu Lys Ala Ala Val Gly Glu Leu Ser Glu Lys Ser Lys  
 65 70 75 80  
 Leu Gln Glu Ile Tyr Gln Glu Leu Thr Gln Leu Lys Ala Ala Val Gly  
 85 90 95

0044553 120159

87

Glu Leu Pro Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr  
 100 105 110  
 Arg Leu Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln  
 115 120 125  
 5 Glu Ile Tyr Gln Glu Leu Thr Arg Leu Lys Ala Ala Val Gly Glu Leu  
 130 135 140  
 Pro Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr Arg Leu  
 145 150 155 160  
 Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln Glu Ile  
 10 165 170 175  
 Tyr Gln Glu Leu Thr Glu Leu Lys Ala Ala Val Gly Glu Leu Pro Glu  
 180 185 190  
 Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr Gln Leu Lys Ala  
 195 200 205  
 15 Ala Val Gly Glu Leu Pro Asp Gln Ser Lys Gln Gln Gln Ile Tyr Gln  
 210 215 220  
 Glu Leu Thr Asp Leu Lys Thr Ala Phe Glu Arg Leu Cys Arg His Cys  
 225 230 235 240  
 Pro Lys Asp Trp Thr Phe Phe Gln Gly Asn Cys Tyr Phe Met Ser Asn  
 20 245 250 255  
 Ser Gln Arg Asn Trp His Asp Ser Val Thr Ala Cys Gln Glu Val Arg  
 260 265 270  
 Ala Gln Leu Val Val Ile Lys Thr Ala Glu Glu Gln Leu Pro Ala Val  
 275 280 285  
 25 Leu Glu Gln Trp Arg Thr Gln Gln  
 290 295

## (2) INFORMATION FOR SEQ ID NO: 4:

- 30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 197  
 (B) TYPE: Amino acid  
 (D) TOPOLOGY: Linear  
 (ii) SEQUENCE KIND: Protein  
 35 (iii) HYPOTHETICAL: No  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Homo sapiens*

04558 12049  
 65T02T 8525460



88

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP01440

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

5

Met Cys Thr Gly Lys Cys Ala Arg Cys Val Gly Leu Ser Leu Ile Thr  
 1 5 10 15

Leu Cys Leu Val Cys Ile Val Ala Asn Ala Leu Leu Leu Val Pro Asn  
 20 25 30

10 Gly Glu Thr Ser Trp Thr Asn Thr Asn His Leu Ser Leu Gln Val Trp  
 35 40 45

Leu Met Gly Gly Phe Ile Gly Gly Gly Leu Met Val Leu Cys Pro Gly  
 50 55 60

Ile Ala Ala Val Arg Ala Gly Gly Lys Gly Cys Cys Gly Ala Gly Cys  
 15 65 70 75 80

Cys Gly Asn Arg Cys Arg Met Leu Arg Ser Val Phe Ser Ser Ala Phe  
 85 90 95

Gly Val Leu Gly Ala Ile Tyr Cys Leu Ser Val Ser Gly Ala Gly Leu  
 100 105 110

20 Arg Asn Gly Pro Arg Cys Leu Met Asn Gly Glu Trp Gly Tyr His Phe  
 115 120 125

Glu Asp Thr Ala Gly Ala Tyr Leu Leu Asn Arg Thr Leu Trp Asp Arg  
 130 135 140

Cys Glu Ala Pro Pro Arg Val Val Pro Trp Asn Val Thr Leu Phe Ser  
 25 145 150 155 160

Leu Leu Val Ala Ala Ser Cys Leu Glu Ile Val Leu Cys Gly Ile Gln  
 165 170 175

Leu Val Asn Ala Thr Ile Gly Val Phe Cys Gly Asp Cys Arg Lys Lys  
 180 185 190

30 Gln Asp Thr Pro His  
 195

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 221

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

09445550120199  
 8525460

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

5 (B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP01526

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

10 Met Glu Ala Gly Gly Phe Leu Asp Ser Leu Ile Tyr Gly Ala Cys Val  
1 5 10 15  
Val Phe Thr Leu Gly Met Phe Ser Ala Gly Leu Ser Asp Leu Arg His  
20 25 30  
Met Arg Met Thr Arg Ser Val Asp Asn Val Gln Phe Leu Pro Phe Leu  
15 35 40 45  
Thr Thr Glu Val Asn Asn Leu Gly Trp Leu Ser Tyr Gly Ala Leu Lys  
50 55 60  
Gly Asp Gly Ile Leu Ile Val Val Asn Thr Val Gly Ala Ala Leu Gln  
65 70 75 80  
20 Thr Leu Tyr Ile Leu Ala Tyr Leu His Tyr Cys Pro Arg Lys Arg Val  
85 90 95  
Val Leu Leu Gln Thr Ala Thr Leu Leu Gly Val Leu Leu Leu Gly Tyr  
100 105 110  
Gly Tyr Phe Trp Leu Leu Val Pro Asn Pro Glu Ala Arg Leu Gln Gln  
25 115 120 125  
Leu Gly Leu Phe Cys Ser Val Phe Thr Ile Ser Met Tyr Leu Ser Pro  
130 135 140  
Leu Ala Asp Leu Ala Lys Val Ile Gln Thr Lys Ser Thr Gln Cys Leu  
145 150 155 160  
30 Ser Tyr Pro Leu Thr Ile Ala Thr Leu Leu Thr Ser Ala Ser Trp Cys  
165 170 175  
Leu Tyr Gly Phe Arg Leu Arg Asp Pro Tyr Ile Met Val Ser Asn Phe  
180 185 190  
Pro Gly Ile Val Thr Ser Phe Ile Arg Phe Trp Leu Phe Trp Lys Tyr  
35 195 200 205  
Pro Gln Glu Gln Asp Arg Asn Tyr Trp Leu Leu Gln Thr  
210 215 220

0944558-12049  
667027-8525460

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10230

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ser Asp Ile Gly Asp Trp Phe Arg Ser Ile Pro Ala Ile Thr Arg  
 1 5 10 15  
 Tyr Trp Phe Ala Ala Thr Val Ala Val Pro Leu Val Gly Lys Leu Gly  
 20 25 30  
 Leu Ile Ser Pro Ala Tyr Leu Phe Leu Trp Pro Glu Ala Phe Leu Tyr  
 35 40 45  
 Arg Phe Gln Ile Trp Arg Pro Ile Thr Ala Thr Phe Tyr Phe Pro Val  
 50 55 60  
 Gly Pro Gly Thr Gly Phe Leu Tyr Leu Val Asn Leu Tyr Phe Leu Tyr  
 65 70 75 80  
 Gln Tyr Ser Thr Arg Leu Glu Thr Gly Ala Phe Asp Gly Arg Pro Ala  
 85 90 95  
 Asp Tyr Leu Phe Met Leu Leu Phe Asn Trp Ile Cys Ile Val Ile Thr  
 100 105 110  
 Gly Leu Ala Met Asp Met Gln Leu Leu Met Ile Pro Leu Ile Met Ser  
 115 120 125  
 Val Leu Tyr Val Trp Ala Gln Leu Asn Arg Asp Met Ile Val Ser Phe  
 130 135 140  
 Trp Phe Gly Thr Arg Phe Lys Ala Cys Tyr Leu Pro Trp Val Ile Leu  
 145 150 155 160  
 Gly Phe Asn Tyr Ile Ile Gly Gly Ser Val Ile Asn Glu Leu Ile Gly  
 165 170 175  
 Asn Leu Val Gly His Leu Tyr Phe Phe Leu Met Phe Arg Tyr Pro Met

91

180 185 190  
 Asp Leu Gly Gly Arg Asn Phe Leu Ser Thr Pro Gln Phe Leu Tyr Arg  
 195 200 205  
 Trp Leu Pro Ser Arg Arg Gly Gly Val Ser Gly Phe Gly Val Pro Pro  
 5 210 215 220  
 Ala Ser Met Arg Arg Ala Ala Asp Gln Asn Gly Gly Gly Gly Arg His  
 225 230 235 240  
 Asn Trp Gly Gln Gly Phe Arg Leu Gly Asp Gln  
 245 250

10

## (2) INFORMATION FOR SEQ ID NO: 7:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Epidermoid carcinoma

(C) CELL LINE: KB

(D) CLONE NAME: HP10389

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Ala Thr Pro Gly Pro Val Ile Pro Glu Val Pro Phe Glu Pro Ser  
 1 5 10 15  
 30 Lys Pro Pro Val Ile Glu Gly Leu Ser Pro Thr Val Tyr Arg Asn Pro  
 20 25 30  
 Glu Ser Phe Lys Glu Lys Phe Val Arg Lys Thr Arg Glu Asn Pro Val  
 35 40 45  
 Val Pro Ile Gly Cys Leu Ala Thr Ala Ala Ala Leu Thr Tyr Gly Leu  
 35 50 55 60  
 Tyr Ser Phe His Arg Gly Asn Ser Gln Arg Ser Gln Leu Met Met Arg  
 65 70 75 80  
 Thr Arg Ile Ala Ala Gln Gly Phe Thr Val Ala Ala Ile Leu Leu Gly

044558 120199

92

85

90

95

Leu Ala Val Thr Ala Met Lys Ser Arg Pro  
 100 105

5

## (2) INFORMATION FOR SEQ ID NO: 8:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10408

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Gly Ser Gly Leu Pro Leu Val Leu Leu Leu Thr Leu Leu Gly Ser  
 1 5 10 15  
 Ser His Gly Thr Gly Pro Gly Met Thr Leu Gln Leu Lys Leu Lys Glu  
 20 25 30  
 Ser Phe Leu Thr Asn Ser Ser Tyr Glu Ser Ser Phe Leu Glu Leu Leu  
 35 40 45  
 Glu Lys Leu Cys Leu Leu Leu His Leu Pro Ser Gly Thr Ser Val Thr  
 50 55 60  
 Leu His His Ala Arg Ser Gln His His Val Val Cys Asn Thr  
 65 70 75

## (2) INFORMATION FOR SEQ ID NO: 9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 314

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Stomach cancer  
 (D) CLONE NAME: HP10412

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

10

Met Val Ala Pro Val Trp Tyr Leu Val Ala Ala Ala Leu Leu Val Gly  
 1 5 10 15  
 Phe Ile Leu Phe Leu Thr Arg Ser Arg Gly Arg Ala Ala Ser Ala Gly  
 20 25 30  
 15 Gln Glu Pro Leu His Asn Glu Glu Leu Ala Gly Ala Gly Arg Val Ala  
 35 40 45  
 Gln Pro Gly Pro Leu Glu Pro Glu Glu Pro Arg Ala Gly Gly Arg Pro  
 50 55 60  
 Arg Arg Arg Arg Asp Leu Gly Ser Arg Leu Gln Ala Gln Arg Arg Ala  
 20 65 70 75 80  
 Gln Arg Val Ala Trp Ala Glu Ala Asp Glu Asn Glu Glu Glu Ala Val  
 85 90 95  
 Ile Leu Ala Gln Glu Glu Glu Gly Val Glu Lys Pro Ala Glu Thr His  
 100 105 110  
 25 Leu Ser Gly Lys Ile Gly Ala Lys Lys Leu Arg Lys Leu Glu Glu Lys  
 115 120 125  
 Gln Ala Arg Lys Ala Gln Arg Glu Ala Glu Glu Ala Glu Arg Glu Glu  
 130 135 140  
 Arg Lys Arg Leu Glu Ser Gln Arg Glu Ala Glu Trp Lys Lys Glu Glu  
 30 145 150 155 160  
 Glu Arg Leu Arg Leu Glu Glu Glu Gln Lys Glu Glu Glu Glu Arg Lys  
 165 170 175  
 Ala Arg Glu Glu Gln Ala Gln Arg Glu His Glu Glu Tyr Leu Lys Leu  
 180 185 190  
 35 Lys Glu Ala Phe Val Val Glu Glu Glu Gly Val Gly Glu Thr Met Thr  
 195 200 205  
 Glu Glu Gln Ser Gln Ser Phe Leu Thr Glu Phe Ile Asn Tyr Ile Lys  
 210 215 220

00445258 12049

Gln Ser Lys Val Val Leu Leu Glu Asp Leu Ala Ser Gln Val Gly Leu  
 225 230 235 240  
 Arg Thr Gln Asp Thr Ile Asn Arg Ile Gln Asp Leu Leu Ala Glu Gly  
 245 250 255  
 5 Thr Ile Thr Gly Val Ile Asp Asp Arg Gly Lys Phe Ile Tyr Ile Thr  
 260 265 270  
 Pro Glu Glu Leu Ala Ala Val Ala Asn Phe Ile Arg Gln Arg Gly Arg  
 275 280 285  
 Val Ser Ile Ala Glu Leu Ala Gln Ala Ser Asn Ser Leu Ile Ala Trp  
 10 290 295 300  
 Gly Arg Glu Ser Pro Ala Gln Ala Pro Ala  
 305 310

15 (2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 195

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

20 (ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

25 (B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10413

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

30 Met Ala Ala Glu Asp Val Val Ala Thr Gly Ala Asp Pro Ser Asp Leu  
 1 5 10 15  
 Glu Ser Gly Gly Leu Leu His Glu Ile Phe Thr Ser Pro Leu Asn Leu  
 20 25 30  
 Leu Leu Leu Gly Leu Cys Ile Phe Leu Leu Tyr Lys Ile Val Arg Gly  
 35 35 40 45  
 Asp Gln Pro Ala Ala Ser Gly Asp Ser Asp Asp Asp Glu Pro Pro Pro  
 50 55 60  
 Leu Pro Arg Leu Lys Arg Arg Asp Phe Thr Pro Ala Glu Leu Arg Arg

09445258-120199  
 667027-85264460

65						70						75						80
Phe	Asp	Gly	Val	Gln	Asp	Pro	Arg	Ile	Leu	Met	Ala	Ile	Asn	Gly	Lys			
					85						90						95	
Val	Phe	Asp	Val	Thr	Lys	Gly	Arg	Lys	Phe	Tyr	Gly	Pro	Glu	Gly	Pro			
					100						105						110	
Tyr	Gly	Val	Phe	Ala	Gly	Arg	Asp	Ala	Ser	Arg	Gly	Leu	Ala	Thr	Phe			
					115						120						125	
Cys	Leu	Asp	Lys	Glu	Ala	Leu	Lys	Asp	Glu	Tyr	Asp	Asp	Leu	Ser	Asp			
					130						135						140	
Leu	Thr	Ala	Ala	Gln	Gln	Glu	Thr	Leu	Ser	Asp	Trp	Glu	Ser	Gln	Phe			
					145						150						155	
Thr	Phe	Lys	Tyr	His	His	Val	Gly	Lys	Leu	Leu	Lys	Glu	Gly	Glu	Glu			
					165						170						175	
Pro	Thr	Val	Tyr	Ser	Asp	Glu	Glu	Glu	Pro	Lys	Asp	Glu	Ser	Ala	Arg			
					180						185						190	
Lys	Asn	Asp																
					195													

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: Amino acid

25 (ii) SEQUENCE KIND: Protein

(vi) ORIGINAL SOURCE:

30 (B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10415

35 Met Leu Asp Phe Ala Ile Phe Ala Val Thr Phe Leu Leu Ala Leu Val

**1                      5                      10                      15**

Gly Ala Val Leu Tyr Leu Tyr Pro Ala Ser Arg Gln Ala Ala Gly Ile

20                      25                      30



Pro Gly Ile Thr Pro Thr Glu Glu Lys Asp Gly Asn Leu Pro Asp Ile  
 35 40 45  
 Val Asn Ser Gly Ser Leu His Glu Phe Leu Val Asn Leu His Glu Arg  
 50 55 60  
 5 Tyr Gly Pro Val Val Ser Phe Trp Phe Gly Arg Arg Leu Val Val Ser  
 65 70 75 80  
 Leu Gly Thr Val Asp Val Leu Lys Gln His Ile Asn Pro Asn Lys Thr  
 85 90 95  
 Leu Asp Pro Phe Glu Thr Met Leu Lys Ser Leu Leu Arg Tyr Gln Ser  
 10 100 105 110  
 Gly Gly Gly Ser Val Ser Glu Asn His Met Arg Lys Lys Leu Tyr Glu  
 115 120 125  
 Asn Gly Val Thr Asp Ser Leu Lys Ser Asn Phe Ala Leu Leu Leu Lys  
 130 135 140  
 15 Leu Ser Glu Glu Leu Leu Asp Lys Trp Leu Ser Tyr Pro Glu Thr Gln  
 145 150 155 160  
 His Val Pro Leu Ser Gln His Met Leu Gly Phe Ala Met Lys Ser Val  
 165 170 175  
 Thr Gln Met Val Met Gly Ser Thr Phe Glu Asp Asp Gln Glu Val Ile  
 20 180 185 190  
 Arg Phe Gln Lys Asn His Gly Thr Val Trp Ser Glu Ile Gly Lys Gly  
 195 200 205  
 Phe Leu Asp Gly Ser Leu Asp Lys Asn Met Thr Arg Lys Lys Gln Tyr  
 210 215 220  
 25 Glu Asp Ala Leu Met Gln Leu Glu Ser Val Leu Arg Asn Ile Ile Lys  
 225 230 235 240  
 Glu Arg Lys Gly Arg Asn Phe Ser Gln His Ile Phe Ile Asp Ser Leu  
 245 250 255  
 Val Gln Gly Asn Leu Asn Asp Gln Gln Ile Leu Glu Asp Ser Met Ile  
 30 260 265 270  
 Phe Ser Leu Ala Ser Cys Ile Ile Thr Ala Lys Leu Cys Thr Trp Ala  
 275 280 285  
 Ile Cys Phe Leu Thr Thr Ser Glu Glu Val Gln Lys Lys Leu Tyr Glu  
 290 295 300  
 35 Glu Ile Asn Gln Val Phe Gly Asn Gly Pro Val Thr Pro Glu Lys Ile  
 305 310 315 320  
 Glu Gln Leu Arg Tyr Cys Gln His Val Leu Cys Glu Thr Val Arg Thr  
 325 330 335

00445558 120199  
 667027 8524460

97

Ala Lys Leu Thr Pro Val Ser Ala Gln Leu Gln Asp Ile Glu Gly Lys  
 340 345 350  
 Ile Asp Arg Phe Ile Ile Pro Arg Glu Thr Leu Val Leu Tyr Ala Leu  
 355 360 365  
 5 Gly Val Val Leu Gln Asp Pro Asn Thr Trp Pro Ser Pro His Lys Phe  
 370 375 380  
 Asp Pro Asp Arg Phe Asp Asp Glu Leu Val Met Lys Thr Phe Ser Ser  
 385 390 395 400  
 Leu Gly Phe Ser Gly Thr Gln Glu Cys Pro Glu Leu Arg Phe Ala Tyr  
 10 405 410 415  
 Met Val Thr Thr Val Leu Leu Ser Val Leu Val Lys Arg Leu His Leu  
 420 425 430  
 Leu Ser Val Glu Gly Gln Val Ile Glu Thr Lys Tyr Glu Leu Val Thr  
 435 440 445  
 15 Ser Ser Arg Glu Glu Ala Trp Ile Thr Val Ser Lys Arg Tyr  
 450 455 460

## (2) INFORMATION FOR SEQ ID NO: 12:

## 20 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 247

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

25 (iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

30 (D) CLONE NAME: HP10419

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Met Gly Ala Ala Val Phe Phe Gly Cys Thr Phe Val Ala Phe Gly Pro  
 35 1 5 10 15  
 Ala Phe Ala Leu Phe Leu Ile Thr Val Ala Gly Asp Pro Leu Arg Val  
 20 25 30  
 Ile Ile Leu Val Ala Gly Ala Phe Phe Trp Leu Val Ser Leu Leu Leu

094453-12049 66T02T 8524460

98

35 40 45  
 Ala Ser Val Val Trp Phe Ile Leu Val His Val Thr Asp Arg Ser Asp  
 50 55 60  
 Ala Arg Leu Gln Tyr Gly Leu Leu Ile Phe Gly Ala Ala Val Ser Val  
 5 65 70 75 80  
 Leu Leu Gln Glu Val Phe Arg Phe Ala Tyr Tyr Lys Leu Leu Lys Lys  
 85 90 95  
 Ala Asp Glu Gly Leu Ala Ser Leu Ser Glu Asp Gly Arg Ser Pro Ile  
 100 105 110  
 10 Ser Ile Arg Gln Met Ala Tyr Val Ser Gly Leu Ser Phe Gly Ile Ile  
 115 120 125  
 Ser Gly Val Phe Ser Val Ile Asn Ile Leu Ala Asp Ala Leu Gly Pro  
 130 135 140  
 Gly Val Val Gly Ile His Gly Asp Ser Pro Tyr Tyr Phe Leu Thr Ser  
 15 145 150 155 160  
 Ala Phe Leu Thr Ala Ala Ile Ile Leu Leu His Thr Phe Trp Gly Val  
 165 170 175  
 Val Phe Phe Asp Ala Cys Glu Arg Arg Arg Tyr Trp Ala Leu Gly Leu  
 180 185 190  
 20 Val Val Gly Ser His Leu Leu Thr Ser Gly Leu Thr Phe Leu Asn Pro  
 195 200 205  
 Trp Tyr Glu Ala Ser Leu Leu Pro Ile Tyr Ala Val Thr Val Ser Met  
 210 215 220  
 Gly Leu Trp Ala Phe Ile Thr Ala Gly Gly Ser Leu Arg Ser Ile Gln  
 25 225 230 235 240  
 Arg Ser Leu Leu Cys Lys Asp  
 245

30 (2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

35 (ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

05445258-120199

99

(A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP10424

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Met Asn Phe Tyr Leu Leu Leu Ala Ser Ser Ile Leu Cys Ala Leu Ile  
1 5 10 15  
Val Phe Trp Lys Tyr Arg Arg Phe Gln Arg Asn Thr Gly Glu Met Ser  
10 20 25 30  
Ser Asn Ser Thr Ala Leu Ala Leu Val Arg Pro Ser Ser Ser Gly Leu  
35 40 45  
Ile Asn Ser Asn Thr Asp Asn Asn Leu Ala Val Tyr Asp Leu Ser Arg  
50 55 60  
15 Asp Ile Leu Asn Asn Phe Pro His Ser Ile Ala Arg Gln Lys Arg Ile  
65 70 75 80  
Leu Val Asn Leu Ser Met Val Glu Asn Lys Leu Val Glu Leu Glu His  
85 90 95  
Thr Leu Leu Ser Lys Gly Phe Arg Gly Ala Ser Pro His Arg Lys Ser  
20 100 105 110  
Thr

## (2) INFORMATION FOR SEQ ID NO: 14:

## 25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 365  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

30 (iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Epidermoid carcinoma  
35 (C) CELL LINE: KB  
(D) CLONE NAME: HP10428

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

100

Met Gly Arg Trp Ala Leu Asp Val Ala Phe Leu Trp Lys Ala Val Leu  
 1 5 10 15  
 Thr Leu Gly Leu Val Leu Leu Tyr Tyr Cys Phe Ser Ile Gly Ile Thr  
 20 25 30  
 5 Phe Tyr Asn Lys Trp Leu Thr Lys Ser Phe His Phe Pro Leu Phe Met  
 35 40 45  
 Thr Met Leu His Leu Ala Val Ile Phe Leu Phe Ser Ala Leu Ser Arg  
 50 55 60  
 Ala Leu Val Gln Cys Ser Ser His Arg Ala Arg Val Val Leu Ser Trp  
 10 65 70 75 80  
 Ala Asp Tyr Leu Arg Arg Val Ala Pro Thr Ala Leu Ala Thr Ala Leu  
 85 90 95  
 Asp Val Gly Leu Ser Asn Trp Ser Phe Leu Tyr Val Thr Val Ser Leu  
 100 105 110  
 15 Tyr Thr Met Thr Lys Ser Ser Ala Val Leu Phe Ile Leu Ile Phe Ser  
 115 120 125  
 Leu Ile Phe Lys Leu Glu Glu Leu Arg Ala Ala Leu Val Leu Val Val  
 130 135 140  
 Leu Leu Ile Ala Gly Gly Leu Phe Met Phe Thr Tyr Lys Ser Thr Gln  
 20 145 150 155 160  
 Phe Asn Val Glu Gly Phe Ala Leu Val Leu Gly Ala Ser Phe Ile Gly  
 165 170 175  
 Gly Ile Arg Trp Thr Leu Thr Gln Met Leu Leu Gln Lys Ala Glu Leu  
 180 185 190  
 25 Gly Leu Gln Asn Pro Ile Asp Thr Met Phe His Leu Gln Pro Leu Met  
 195 200 205  
 Phe Leu Gly Leu Phe Pro Leu Phe Ala Val Phe Glu Gly Leu His Leu  
 210 215 220  
 Ser Thr Ser Glu Lys Ile Phe Arg Phe Gln Asp Thr Gly Leu Leu Leu  
 30 225 230 235 240  
 Arg Val Leu Gly Ser Leu Phe Leu Gly Gly Ile Leu Ala Phe Gly Leu  
 245 250 255  
 Gly Phe Ser Glu Phe Leu Leu Val Ser Arg Thr Ser Ser Leu Thr Leu  
 260 265 270  
 35 Ser Ile Ala Gly Ile Phe Lys Glu Val Cys Thr Leu Leu Leu Ala Ala  
 275 280 285  
 His Leu Leu Gly Asp Gln Ile Ser Leu Leu Asn Trp Leu Gly Phe Ala  
 290 295 300

00445258 120199

101

Leu Cys Leu Ser Gly Ile Ser Leu His Val Ala Leu Lys Ala Leu His  
 305 310 315 320  
 Ser Arg Gly Asp Gly Gly Pro Lys Ala Leu Lys Gly Leu Gly Ser Ser  
 325 330 335  
 5 Pro Asp Leu Glu Leu Leu Leu Arg Ser Ser Gln Arg Glu Glu Gly Asp  
 340 345 350  
 Asn Glu Glu Glu Glu Tyr Phe Val Ala Gln Gly Gln Gln  
 355 360 365

10

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 226

(B) TYPE: Amino acid

15

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

20

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

25

Met Pro Thr Thr Lys Lys Thr Leu Met Phe Leu Ser Ser Phe Phe Thr  
 1 5 10 15  
 Ser Leu Gly Ser Phe Ile Val Ile Cys Ser Ile Leu Gly Thr Gln Ala  
 20 25 30  
 30 Trp Ile Thr Ser Thr Ile Ala Val Arg Asp Ser Ala Ser Asn Gly Ser  
 35 40 45  
 Ile Phe Ile Thr Tyr Gly Leu Phe Arg Gly Glu Ser Ser Glu Glu Leu  
 50 55 60  
 Ser His Gly Leu Ala Glu Pro Lys Lys Lys Phe Ala Val Leu Glu Ile  
 35 65 70 75 80  
 Leu Asn Asn Ser Ser Gln Lys Thr Leu His Ser Val Thr Ile Leu Phe  
 85 90 95  
 Leu Val Leu Ser Leu Ile Thr Ser Leu Leu Ser Ser Gly Phe Thr Phe

044525-120190

102

100 105 110  
 Tyr Asn Ser Ile Ser Asn Pro Tyr Gln Thr Phe Leu Gly Pro Thr Gly  
 115 120 125  
 Val Tyr Thr Trp Asn Gly Leu Gly Ala Ser Phe Val Phe Val Thr Met  
 5 130 135 140  
 Ile Leu Phe Val Ala Asn Thr Gln Ser Asn Gln Leu Ser Glu Glu Leu  
 145 150 155 160  
 Phe Gln Met Leu Tyr Pro Ala Thr Thr Ser Lys Gly Thr Thr His Ser  
 165 170 175  
 10 Tyr Gly Tyr Ser Phe Trp Leu Ile Leu Leu Val Ile Leu Leu Asn Ile  
 180 185 190  
 Val Thr Val Thr Ile Ile Ile Phe Tyr Gln Lys Ala Arg Tyr Gln Arg  
 195 200 205  
 Lys Gln Glu Gln Arg Lys Pro Met Glu Tyr Ala Pro Arg Asp Gly Ile  
 15 210 215 220  
 Leu Phe  
 225

20 (2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 129

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

25 (ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

30 (B) CELL KIND: Liver

(D) CLONE NAME: HP10432

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

35 Met Ala Arg Gly Ser Leu Arg Arg Leu Leu Arg Leu Leu Val Leu Gly  
 1 5 10 15  
 Leu Trp Leu Ala Leu Leu Arg Ser Val Ala Gly Glu Gln Ala Pro Gly  
 20 25 30

044558-12019

103

Thr Ala Pro Cys Ser Arg Gly Ser Ser Trp Ser Ala Asp Leu Asp Lys  
 35 40 45  
 Cys Met Asp Cys Ala Ser Cys Arg Ala Arg Pro His Ser Asp Phe Cys  
 50 55 60  
 5 Leu Gly Cys Ala Ala Ala Pro Pro Ala Pro Phe Arg Leu Leu Trp Pro  
 65 70 75 80  
 Ile Leu Gly Gly Ala Leu Ser Leu Thr Phe Val Leu Gly Leu Leu Ser  
 85 90 95  
 Gly Phe Leu Val Trp Arg Arg Cys Arg Arg Arg Glu Lys Phe Thr Thr  
 10 100 105 110  
 Pro Ile Glu Glu Thr Gly Gly Glu Gly Cys Pro Ala Val Ala Leu Ile  
 115 120 125  
 Gln

15

## (2) INFORMATION FOR SEQ ID NO: 17:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 163

20

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

25

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Liver

(D) CLONE NAME: HP10433

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Met Arg Arg Leu Leu Ile Pro Leu Ala Leu Trp Leu Gly Ala Val Gly  
 1 5 10 15  
 Val Gly Val Ala Glu Leu Thr Glu Ala Gln Arg Arg Gly Leu Gln Val  
 35 20 25 30  
 Ala Leu Glu Glu Phe His Lys His Pro Pro Val Gln Trp Ala Phe Gln  
 35 40 45  
 Glu Thr Ser Val Glu Ser Ala Val Asp Thr Pro Phe Pro Ala Gly Ile

0445258 120199



104

50                      55                      60  
 Phe Val Arg Leu Glu Phe Lys Leu Gln Gln Thr Ser Cys Arg Lys Arg  
 65                      70                      75                      80  
 Asp Trp Lys Lys Pro Glu Cys Lys Val Arg Pro Asn Gly Arg Lys Arg  
 5                      85                      90                      95  
 Lys Cys Leu Ala Cys Ile Lys Leu Gly Ser Glu Asp Lys Val Leu Gly  
 100                      105                      110  
 Arg Leu Val His Cys Pro Ile Glu Thr Gln Val Leu Arg Glu Ala Glu  
 115                      120                      125  
 10 Glu His Gln Glu Thr Gln Cys Leu Arg Val Gln Arg Ala Gly Glu Asp  
 130                      135                      140  
 Pro His Ser Phe Tyr Phe Pro Gly Gln Phe Ala Phe Ser Lys Ala Leu  
 145                      150                      155                      160  
 Pro Arg Ser

## (2) INFORMATION FOR SEQ ID NO: 18:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 193

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: Protein

(iii) HYPOTHETICAL: No

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10480

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Ile Arg Cys Gly Leu Ala Cys Glu Arg Cys Arg Trp Ile Leu Pro  
 1                      5                      10                      15  
 Leu Leu Leu Leu Ser Ala Ile Ala Phe Asp Ile Ile Ala Leu Ala Gly  
 35                      20                      25                      30  
 Arg Gly Trp Leu Gln Ser Ser Asp His Gly Gln Thr Ser Ser Leu Trp  
 35                      40                      45  
 Trp Lys Cys Ser Gln Glu Gly Gly Gly Ser Gly Ser Tyr Glu Glu Gly

00445558 120160  
 00445558 120160

[illegible]

ATGGGTCTGC	TCCTTCCCCT	GGCACTCTGC	ATCCTAGTCC	TGTGCTGCGG	AGCAATGTCT	60
CCACCCCAGC	TGGCCCTCAA	CCCCTCGGCT	CTGCTCTCCC	GGGGCTGCAA	TGACTCCGAT	120
GTGCTGGCAG	TTGCAGGCTT	TGCCCTGCGG	GATATTAACA	AAGACAGAAA	GGATGGCTAT	180

106

GTGCTGAGAC TCAACCGAGT GAACGACGCC CAGGAATACA GACGGGGTGG CCTGGGATCT 240  
CTGTTCTATC TTACACTGGA TGTGCTAGAG ACTGACTGCC ATGTGCTCAG AAAGAAGGCA 300  
TGGCAAGACT GTGGAATGAG GATATTTTTT GAATCAGTTT ATGGTCAATG CAAAGCAATA 360  
TTTTATATGA ACAACCCAAG TAGAGTTCTC TATTTAGCTG CTTATAACTG TACTCTTCGC 420  
5 CCAGTTTCAA AAAAAAAGAT TTACATGACG TGCCCTGACT GCCCAAGCTC CATACCCACT 480  
GACTCTTCCA ATCACCAAGT GCTGGAGGCT GCCACCGAGT CTCTTGCGAA ATACAACAAT 540  
GAGAACACAT CCAAGCAGTA TTCTCTCTTC AAAGTCACCA GGGCTTCTAG CCAGTGGGTG 600  
GTCGGCCCTT CTTACTTTGT GGAATACTTA ATTAAAGAAT CACCATGTAC TAAATCCCAG 660  
GCCAGCAGCT GTTCACTTCA GTCCTCCGAC TCTGTGCCTG TTGGTCTTTG CAAAGGTTCT 720  
10 CTGACTCGAA CACACTGGGA AAAGTTTGTC TCTGTGACTT GTGACTTCTT TGAATCACAG 780  
GCTCCAGCCA CTGGAAGTGA AAACCTCTGCT GTTAACCAGA AACCTACAAA CCTTCCCAAG 840  
GTGGAAGAAT CCCAGCAGAA AAACACCCCC CCAACAGACT CCCCCTCCAA AGCTGGGCCA 900  
AGAGGATCTG TCCAATATCT TCCTGACTTG GATGATAAAA ATTCCCAGGA AAAGGGCCCT 960  
CAGGAGGCCT TTCCTGTGCA TCTGGACCTA ACCACGAATC CCCAGGGAGA AACCCTGGAT 1020  
15 ATTTCTTCC TCTTCCTGGA GCCTATGGAG GAGAAGCTGG TTGTCCTGCC TTTCCCCAAA 1080  
GAAAAAGCAC GCACTGCTGA GTGCCAGGG CCAGCCCAGA ATGCCAGCCC TCTTGTCTT 1140  
CCGCCA 1146

20 (2) INFORMATION FOR SEQ ID NO: 20:  
(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 951  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
25 (D) TOPOLOGY: Linear  
(ii) SEQUENCE KIND: cDNA to mRNA  
(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Homo sapiens*  
30 (B) CELL KIND: Liver  
(D) CLONE NAME: HP01299

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

35 ATGTGGCTCT ACCTGGCGGC CTCGTGGGC CTGTACTACC TTCTGCACTG GTACCGGGAG 60  
AGGCAGGTGG TGAGCCACCT CCAAGACAAG TATGTCTTTA TCACGGGCTG TGA CTGGGC 120  
TTTGGGAACC TGCTGGCCAG ACAGCTGGAT GCACGAGGCT TGAGAGTGCT GGCTGCGTGT 180  
CTGACGGAGA AGGGGGCCGA GCAGCTGAGG GGCCAGACGT CTGACAGGCT GGAGACGGTG 240

107

ACCCTGGATG TTACCAAGAT GGAGAGCATC GCTGCAGCTA CTCAGTGGGT GAAGGAGCAT 300  
GTGGGGGACA GAGGACTCTG GGGACTGGTG AACAAATGCAG GCATTCTTAC ACCAATTACC 360  
TTATGTGAGT GGCTGAACAC TGAGGACTCT ATGAATATGC TCAAAGTGAA CCTCATTGGT 420  
GTGATCCAGG TGACCTTGAG CATGCTTCCT TTGGTGAGGA GAGCACGGGG AAGAATTGTC 480  
5 AATGTCTCCA GCATTCTGGG AAGAGTTGCT TTCTTTGTAG GAGGCTACTG TGTCTCCAAG 540  
TATGGAGTGG AAGCCTTTTC AGATATTCTG AGGCGTGAGA TTCAACATTT TGGGGTGAAA 600  
ATCAGCATAG TTGAACCTGG CTACTTCAGA ACGGGAATGA CAAACATGAC ACAGTCCTTA 660  
GAGCGAATGA AGCAAAGTTG GAAAGAAGCC CCCAAGCATA TTAAGGAGAC CTATGGACAG 720  
CAGTATTTTG ATGCCCTTTA CAATATCATG AAGGAAGGGC TGTTGAATTG TAGCACAAAC 780  
10 CTGAACCTGG TCACTGACTG CATGGAACAT GCTCTGACAT CGGTGCATCC GCGAACTCGA 840  
TATTCAGCTG GCTGGGATGC TAAATTTTTC TTCATCCCTC TATCTTATTT ACCTACATCA 900  
CTGGCAGACT ACATTTTGAC TAGATCTTGG CCCAAACCAG CCCAGGCAGT C 951

15 (2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 888

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

20 (D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

25 (B) CELL KIND: Liver

(D) CLONE NAME: HP01347

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

30 ATGAGTGA CT CCAAGGAACC AAGGGTGCAG CAGCTGGGCC TCCTGGGGTG TCTTGGCCAT 60  
GGCGCCCTGG TGCTGCAACT CCTCTCCTTC ATGCTCTTGG CTGGGGTCCT GGTGGCCATC 120  
CTTGTCCAAG TGTCCAAGGT CCCAGCTCC CTAAGTCAGG AACAAATCCGA GCAAGACGCA 180  
ATCTACCAGA ACCTGACCCA GCTTAAAGCT GCAGTGGGTG AGCTCTCAGA GAAATCCAAG 240  
CTGCAGGAGA TCTACCAGGA GCTGACCCAG CTGAAGGCTG CAGTGGGTGA GTTGCCAGAG 300  
35 AAATCCAAGC TGCAGGAGAT CTACCAGGAG CTGACCCGGC TGAAGGCTGC AGTGGGTGAG 360  
TTGCCAGAGA AATCCAAGCT GCAGGAGATC TACCAGGAGC TGACCCGGCT GAAGGCTGCA 420  
GTGGGTGAGT TGCCAGAGAA ATCCAAGCTG CAGGAGATCT ACCAGGAGCT GACCCGGCTG 480  
AAGGCTGCAG TGGGTGAGTT GCCAGAGAAA TCCAAGCTGC AGGAGATCTA CCAGGAGCTG 540

108

ACGGAGCTGA AGGCTGCAGT GGGTGAGTTG CCAGAGAAAT CCAAGCTGCA GGAGATCTAC 600  
CAGGAGCTGA CCCAGCTGAA GGCTGCAGTG GGTGAGTTGC CAGACCAGTC CAAGCAGCAG 660  
CAAATCTATC AAGAACTGAC CGATTTGAAG ACTGCATTTG AACGCCTGTG CCGCCACTGT 720  
CCCAAGGACT GGACATTCTT CCAAGGAAAC TGTTACTTCA TGTCTAACTC CCAGCGGAAC 780  
5 TGGCACGACT CCGTCACCGC CTGCCAGGAA GTGAGGGCCC AGCTCGTCGT AATCAAAACT 840  
GCTGAGGAGC AGCTTCCAGC GGTACTGGAA CAGTGGAGAA CCCAACAA 888

## (2) INFORMATION FOR SEQ ID NO: 22:

## 10 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 591

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## 15 (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

## 20 (D) CLONE NAME: HP01440

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ATGTGTACGG GAAAATGTGC CCGCTGTGTG GGGCTCTCCC TCATTACCCT CTGCCTCGTC 60  
25 TGCATTGTGG CCAACGCCCT CCTGCTGGTA CCTAATGGGG AGACCTCCTG GACCAACACC 120  
AACCATCTCA GCTTGCAAGT CTGGCTCATG GGCGGCTTCA TTGGCGGGGG CCTAATGGTA 180  
CTGTGTCCGG GGATTGCAGC CGTTCGGGCA GGGGGCAAGG GCTGCTGTGG TGCTGGGTGC 240  
TGTGGAAACC GCTGCAGGAT GCTGCGCTCG GTCTTCTCCT CGGCGTTCGG GGTGCTTGGT 300  
GCCATCTACT GCCTCTCGGT GTCTGGAGCT GGGCTCCGAA ATGGACCCAG ATGCTTAATG 360  
30 AACGGCGAGT GGGGCTACCA CTTCGAAGAC ACCGCGGGAG CTTACTTGCT CAACCGCACT 420  
CTATGGGATC GGTGCGAGGC GCCCCCTCGC GTGGTCCCCT GGAATGTGAC GCTCTTCTCG 480  
CTGCTGGTGG CCGCCTCCTG CCTGGAGATA GTACTGTGTG GGATCCAGCT GGTGAACGCG 540  
ACCATTGGTG TCTTCTGCGG CGATTGCAGG AAAAAACAGG ACACCCCTCA C 591

35

## (2) INFORMATION FOR SEQ ID NO: 23:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 663

109

- (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP01526

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATGGAGGCGG GCGGCTTTCT GGACTCGCTC ATTTACGGAG CATGCGTGGT CTTACCCTT 60  
GGCATGTTCT CCGCCGGCCT CTCGGACCTC AGGCACATGC GAATGACCCG GAGTGTGGAC 120  
AACGTCCAGT TCCTGCCCTT TCTCACCACG GAAGTCAACA ACCTGGGCTG GCTGAGTTAT 180  
GGGGCTTTGA AGGGAGACGG GATCCTCATC GTCGTCAACA CAGTGGGTGC TCGCCTTCAG 240  
ACCCTGTATA TCTTGGCATA TCTGCATTAC TGCCCTCGGA AGCGTGTTGT GCTCCTACAG 300  
ACTGCAACCC TGCTAGGGGT CCTTCTCCTG GGTATGGCT ACTTTTGGCT CCTGGTACCC 360  
AACCCTGAGG CCCGGCTTCA GCAGTTGGGC CTCTTCTGCA GTGTCTTCAC CATCAGCATG 420  
TACCTCTCAC CACTGGCTGA CTTGGCTAAG GTGATTCAAA CTAAATCAAC CCAATGTCTC 480  
TCCTACCCAC TCACCATTGC TACCCTTCTC ACCTCTGCCT CCTGGTGCCT CTATGGGTTT 540  
CGACTCAGAG ATCCCTATAT CATGGTGTCC AACTTTCCAG GAATCGTCAC CAGCTTTATC 600  
CGCTTCTGGC TTTTCTGGAA GTACCCCCAG GAGCAAGACA GGAAGTACTG GCTCCTGCAA 660  
ACC 663

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 753  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP10230

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

ATGTCGGACA TCGGAGACTG GTTCAGGAGC ATCCCGGCCA TCACGCGCTA TTGGTTCGCC 60  
GCCACCGTCG CCGTGCCCTT GGTCCGCAAA CTCGGCCTCA TCAGCCCGGC CTACCTCTTC 120  
5 CTCTGGCCCG AAGCCTTCCT TTATCGCTTT CAGATTGGA GGCCAATCAC TGCCACCTTT 180  
TATTTCCCTG TGGGTCCAGG AACTGGATTT CTTTATTTGG TCAATTTATA TTTCTTATAT 240  
CAGTATTCTA CGCGACTTGA AACAGGAGCT TTTGATGGGA GGCCAGCAGA CTATTTATTC 300  
ATGCTCCTCT TTAAGTGGAT TTGCATCGTG ATTACTGGCT TAGCAATGGA TATGCAGTTG 360  
CTGATGATTC CTCTGATCAT GTCAGTACTT TATGTCTGGG CCCAGCTGAA CAGAGACATG 420  
10 ATTGTATCAT TTTGGTTTGG AACACGATTT AAGGCCTGCT ATTTACCCTG GGTTATCCTT 480  
GGATTCAACT ATATCATCGG AGGCTCGGTA ATCAATGAGC TTATTGGAAA TCTGGTTGGA 540  
CATCTTTATT TTTTCCTAAT GTTCAGATAC CCAATGGACT TGGGAGGAAG AAATTTTCTA 600  
TCCACACCTC AGTTTTTGTA CCGCTGGCTG CCCAGTAGGA GAGGAGGAGT ATCAGGATTT 660  
GGTGTGCCCC CTGCTAGCAT GAGGCGAGCT GCTGATCAGA ATGGCGGAGG CGGGAGACAC 720  
15 AACTGGGGCC AGGGCTTTTCG ACTTGAGAC CAG 753

## (2) INFORMATION FOR SEQ ID NO: 25:

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 318  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear  
(ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- 25 (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Epidermoid carcinoma  
(C) CELL LINE: KB  
30 (D) CLONE NAME: HP10389

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

ATGGCGACTC CCGGCCCTGT GATTCCGGAG GTCCCCTTTG AACCATCGAA GCCTCCAGTC 60  
35 ATTGAGGGGC TGAGCCCCAC TGTTTACAGG AATCCAGAGA GTTTCAGGA AAAGTTCGTT 120  
CGCAAGACCC GCGAGAACCC GGTGGTACCC ATAGGTTGCC TGGCCACGGC GGCCGCCCTC 180  
ACCTACGGCC TCTACTCCTT CCACCGGGGC AACAGCCAGC GCTCTCAGCT CATGATGCGC 240  
ACCCGGATCG CCGCCAGGG TTTACGGTC GCAGCCATCT TGCTGGGTCT GGCTGTCACT 300

111

GCTATGAAGT CTCGACCC

318

## (2) INFORMATION FOR SEQ ID NO: 26:

- 5 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 234
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 10 (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- 15 (A) ORGANISM: *Homo sapiens*
- (B) CELL KIND: Stomach cancer
- (D) CLONE NAME: HP10408

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ATGGGGTCTG GGCTGCCCCCT TGTCTCCTC TTGACCCTCC TTGGCAGCTC ACATGGAACA 60

20 GGGCCGGGTA TGACTTTGCA ACTGAAGCTG AAGGAGTCTT TTCTGACAAA TTCCTCCTAT 120

GAGTCCAGCT TCCTGGAATT GCTTGAAAAG CTCTGCCTCC TCCTCCATCT CCCTTCAGGG 180

ACCAGCGTCA CCCTCCACCA TGCAAGATCT CAACACCATG TTGTCTGCAA CACA 234

## 25 (2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 942
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 30 (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- 35 (A) ORGANISM: *Homo sapiens*
- (B) CELL KIND: Stomach cancer
- (D) CLONE NAME: HP10412

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

00445258.120199



112

ATGGTGGCGC CTGTGTGGTA CTTGGTAGCG GCGGCTCTGC TAGTCGGCTT TATCCTCTTC 60  
CTGACTCGCA GCCGGGGCCG GCGGCATCA GCCGGCCAAG AGCCACTGCA CAATGAGGAG 120  
CTGGCAGGAG CAGGCCGGGT GGCCACGCTT GGGCCCCTGG AGCCTGAGGA GCCGAGAGCT 180  
GGAGGCAGGC CTCGGCGCCG GAGGGACCTG GGCAGCCGCC TACAGGCCCCA GCGTCGAGCC 240  
5 CAGCGGGTGG CCTGGGCAGA AGCAGATGAG AACGAGGAGG AAGCTGTCAT CCTAGCCCAG 300  
GAGGAGGAAG GTGTCGAGAA GCCAGCGGAA ACTCACCTGT CGGGGAAAAT TGGAGCTAAG 360  
AAACTGCGGA AGCTGGAGGA GAAACAAGCG CGAAAGGCC AGCGTGAGGC AGAGGAGGCT 420  
GAACGTGAGG AGCGGAAACG ACTCGAGTCC CAGCGCGAAG CTGAGTGGA GAAGGAGGAG 480  
GAGCGGCTTC GCCTGGAGGA GGAGCAGAAG GAGGAGGAGG AGAGGAAGGC CCGCGAGGAG 540  
10 CAGGCCCAGC GGGAGCATGA GGAGTACCTG AACTGAAGG AGGCCTTTGT GGTGGAGGAG 600  
GAAGGCGTAG GAGAGACCAT GACTGAGGAA CAGTCCAGA GCTTCCTGAC AGAGTTCATC 660  
AACTACATCA AGCAGTCCAA GGTGTGCTC TTGGAAGACC TGGCTTCCCA GGTGGGCCTA 720  
CGCACTCAGG ACACCATAAA TCGCATCCAG GACCTGCTGG CTGAGGGGAC TATAACAGGT 780  
GTGATTGACG ACCGGGGCAA GTTCATCTAC ATAACCCAG AGGAACTGGC CGCCGTGGCC 840  
15 AACTTCATCC GACAGCGGGG CCGGGTGTCC ATCGCCGAGC TTGCCCAAGC CAGCAACTCC 900  
CTCATCGCCT GGGGCCGGGA GTCCCTGCC CAAGCCCCAG CC 942

## (2) INFORMATION FOR SEQ ID NO: 28:

- 20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 585  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear  
25 (ii) SEQUENCE KIND: cDNA to mRNA  
(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
30 (D) CLONE NAME: HP10413

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ATGGCTGCCG AGGATGTGGT GGCGACTGGC GCCGACCCAA GCGATCTGGA GAGCGGCGGG 60  
35 CTGCTGCATG AGATTTTCAC GTCGCCGCTC AACCTGCTGC TGCTTGGCCT CTGCATCTTC 120  
CTGCTCTACA AGATCGTGCG CGGGGACCAG CCGGCGGCCA GCGGCGACAG CGACGACGAC 180  
GAGCCGCCCC CTCTGCCCCG CCTCAAGCGG CGCGACTTCA CCCCCGCCGA GCTGCGGCGC 240  
TTCGACGGCG TCCAGGACCC GCGCATACTC ATGGCCATCA ACGCAAGGT GTTCGATGTG 300

ACCAAAGGCC GCAAATTCTA CGGGCCCGAG GGGCCGTATG GGGTCTTTGC TGGAAGAGAT 360  
GCATCCAGGG GCCTTGCCAC ATTTTGCCTG GATAAGGAAG CACTGAAGGA TGAGTACGAT 420  
GACCTTTCTG ACCTCACTGC TGCCCAGCAG GAGACTCTGA GTGACTGGGA GTCTCAGTTC 480  
ACTTTCAAGT ATCATCACGT GGGCAAACCTG CTGAAGGAGG GGGAGGAGCC CACTGTGTAC 540  
5 TCAGATGAGG AAGAACCAAA AGATGAGAGT GCCCGGAAAA ATGAT 585

## (2) INFORMATION FOR SEQ ID NO: 29:

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1386  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear  
15 (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP10415  
20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ATGTTGGACT TCGCGATCTT CGCCGTTACC TTCTTGCTGG CGTTGGTGGG AGCCGTGCTC 60  
TACCTCTATC CGGCTTCCAG ACAAGCTGCA GGAATTCCAG GGATTACTCC AACTGAAGAA 120  
25 AAAGATGGTA ATCTTCCAGA TATTGTGAAT AGTGGAAGTT TGCATGAGTT CCTGGTTAAT 180  
TTGCATGAGA GATATGGGCC TGTGGTCTCC TTCTGGTTTG GCAGGCGCCT CGTGGTTAGT 240  
TTGGGCACTG TTGATGTACT GAAGCAGCAT ATCAATCCCA ATAAGACATT GGACCCTTTT 300  
GAAACCATGC TGAAGTCATT ATTAAGGTAT CAATCTGGTG GTGGCAGTGT GAGTGAAAAC 360  
CACATGAGGA AAAAATTGTA TGAAAATGGT GTGACTGATT CTCTGAAGAG TAACTTTGCC 420  
30 CTCCTCCTAA AGCTTTCAGA AGAATTATTA GATAAATGGC TCTCCTACCC AGAGACCCAG 480  
CACGTGCCCC TCAGCCAGCA TATGCTTGGT TTTGCTATGA AGTCTGTTAC ACAGATGGTA 540  
ATGGGTAGTA CATTTGAAGA TGATCAGGAA GTCATTGCTT TCCAGAAGAA TCATGGCACA 600  
GTTTGGTCTG AGATTGGAAG AGGCTTTCTA GATGGGTCAC TTGATAAAAA CATGACTCGG 660  
AAAAACAAT ATGAAGATGC CCTCATGCAA CTGGAGTCTG TTTTAAGGAA CATCATAAAA 720  
35 GAACGAAAAG GAAGGAACTT CAGTCAACAT ATTTTCATTG ACTCCTTAGT ACAAGGGAAC 780  
CTTAATGACC AACAGATCCT AGAAGACAGT ATGATATTTT CTCTGGCCAG TTGCATAATA 840  
ACTGCAAAAT TGTGTACCTG GGCAATCTGT TTTTAAACCA CCTCTGAAGA AGTTCAAAAA 900  
AAATTATATG AAGAGATAAA CCAAGTTTTT GGAAATGGTC CTGTTACTCC AGAGAAAATT 960

114

GAGCAGCTCA GATATTGTCA GCATGTGCTT TGTGAAACTG TTCGAACTGC CAAACTGACT 1020  
CCAGTTTCTG CCCAGCTTCA AGATATTGAA GGAAAAATTG ACCGATTTAT TATTCCTAGA 1080  
GAGACCCTCG TCCTTTATGC CCTTGGTGTG GACTTCAGG ATCCTAATAC TTGGCCATCT 1140  
CCACACAAGT TTGATCCAGA TCGGTTTGAT GATGAATTAG TAATGAAAAC TTTTTCCTCA 1200  
5 CTTGGATTCT CAGGCACACA GGAGTGTCCA GAGTTGAGGT TTGCATATAT GGTGACCACA 1260  
GTACTTCTTA GTGTATTGGT GAAGAGACTG CACCTACTTT CTGTGGAGGG ACAGGTTATT 1320  
GAAACAAAGT ATGAACTGGT AACATCATCA AGGGAAGAAG CTTGGATCAC TGTCTCAAAG 1380  
AGATAT 1386

10

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 741

(B) TYPE: Nucleic acid

15

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

20

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10419

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

25

ATGGGGGCTG CGGTGTTTTT CGGCTGCACT TTCGTGCGCT TCGGCCCGGC CTTGCGGCTT 60  
TTCTTGATCA CTGTGGCTGG GGACCCGCTT CGCGTTATCA TCCTGGTCGC AGGGGCATTT 120  
TTCTGGCTGG TCTCCCTGCT CCTGGCCTCT GTGGTCTGGT TCATCTTGGT CCATGTGACC 180  
GACCGGTCAG ATGCCCGGCT CCAGTACGGC CTCCTGATTT TTGGTGCTGC TGTCTCTGTC 240  
30 CTTCTACAGG AGGTGTTCCG CTTTGCCTAC TACAAGCTGC TTAAGAAGGC AGATGAGGGG 300  
TTAGCATCGC TGAGTGAGGA CGGAAGATCA CCCATCTCCA TCCGCCAGAT GGCCTATGTT 360  
TCTGGTCTCT CCTTCGGTAT CATCAGTGGT GTCTTCTCTG TTATCAATAT TTTGGCTGAT 420  
GCACTTGGGC CAGGTGTGGT TGGGATCCAT GGAGACTCAC CCTATTACTT CCTGACTTCA 480  
GCCTTTCTGA CAGCAGCCAT TATCCTGCTC CATACTTTT GGGGAGTTGT GTTCTTTGAT 540  
35 GCCTGTGAGA GGAGACGGA CTGGGCTTTG GGCCTGGTGG TTGGGAGTCA CCTACTGACA 600  
TCGGGACTGA CATTCCTGAA CCCCTGGTAT GAGGCCAGCC TGCTGCCCCAT CTATGCAGTC 660  
ACTGTTTCCA TGGGGCTCTG GGCCTTCATC ACAGCTGGAG GGTCCCTCCG AAGTATTCAG 720  
CGCAGCCTCT TGTGTAAGGA C 741

## (2) INFORMATION FOR SEQ ID NO: 31:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10424

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

ATGAACTTCT ATTTACTCCT AGCGAGCAGC ATTCTGTGTG CCTTGATTGT CTTCTGGAAA 60  
TATCGCCGCT TTCAGAGAAA CACTGGCGAA ATGTCATCAA ATTCAACTGC TCTTGCACTA 120  
GTGAGACCCT CTTCTTCTGG GTTAATTAAC AGCAATACAG ACAACAATCT TGCAGTCTAC 180  
GACCTCTCTC GGGATATTTT AAATAATTTC CCACACTCAA TAGCCAGGCA GAAGCGAATA 240  
TTGGTAAACC TCAGTATGGT GGAAAACAAG CTGGTTGAAC TGGAACATAC TCTACTTAGC 300  
AAGGGTTTCA GAGGTGCATC ACCTCACCGG AAATCCACC 339

## (2) INFORMATION FOR SEQ ID NO: 32:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1095

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Epidermoid carcinoma

(C) CELL LINE: KB

(D) CLONE NAME: HP10428

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

ATGGGGAGGT GGGCCCTCGA TGTGGCCTTT TTGTGGAAGG CGGTGTTGAC CCTGGGGCTG 60  
GTGCTTCTCT ACTACTGCTT CTCCATCGGC ATCACCTTCT ACAACAAGTG GCTGACAAAG 120  
5 AGCTTCCATT TCCCCCTCTT CATGACGATG CTGCACCTGG CCGTGATCTT CCTCTTCTCC 180  
GCCCTGTCCA GGGCGCTGGT TCAGTGCTCC AGCCACAGGG CCCGTGTGGT GCTGAGCTGG 240  
GCCGACTACC TCAGAAGAGT GGCTCCCACA GCTCTGGCGA CGGCGCTTGA CGTGGGCTTG 300  
TCCAAGTGA GCTTCCTGTA TGTCACCGTC TCGCTGTACA CAATGACCAA ATCCTCAGCT 360  
GTCCTCTTCA TCTTGATCTT CTCTCTGATC TTCAAGCTGG AGGAGCTGCG CGCGGCACTG 420  
10 GTCCTGGTGG TCCTCCTCAT CGCCGGGGGT CTCTTCATGT TCACCTACAA GTCCACACAG 480  
TTCAACGTGG AGGGCTTCGC CTTGGTGCTG GGGGCCTCGT TCATCGGTGG CATTGCTGG 540  
ACCCTCACCC AGATGCTCCT GCAGAAGGCT GAACTCGGCC TCCAGAATCC CATCGACACC 600  
ATGTTCCACC TGCAGCCACT CATGTTCTTG GGGCTCTTCC CTCTCTTTGC TGTATTTGAA 660  
GGTCTCCATT TGTCCACATC TGAGAAAATC TTCCGTTTCC AGGACACAGG GCTGCTCCTG 720  
15 CGGGTACTTG GGAGCCTCTT CCTTGGCGGG ATTCTCGCCT TTGGTTTGGG CTTCTCTGAG 780  
TTCCTCCTGG TCTCCAGAAC CTCCAGCCTC ACTCTCTCCA TTGCCGGCAT TTTTAAGGAA 840  
GTCTGCACTT TGCTGTTGGC AGCTCATCTG CTGGGCGATC AGATCAGCCT CCTGAACTGG 900  
CTGGGCTTCG CCCTCTGCCT CTCGGGAATA TCCCTCCACG TTGCCCTCAA AGCCCTGCAT 960  
TCCAGAGGTG ATGGTGGCCC CAAGGCCTTG AAGGGGCTGG GCTCCAGCCC CGACCTGGAG 1020  
20 CTGCTGCTCC GGAGCAGCCA GCGGGAGGAA GGTGACAATG AGGAGGAGGA GTACTTTGTG 1080  
GCCCAGGGGC AGCAG 1095

(2) INFORMATION FOR SEQ ID NO: 33:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 678

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

30 (ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

35 (D) CLONE NAME: HP10429

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

117

ATGCCTACCA CAAAGAAGAC ATTGATGTTT TTATCAAGCT TTTTCACCAG CCTTGGGTCC 60  
TTCATTGTAA TTTGCTCTAT TCTTGGGACA CAAGCATGGA TCACCAGTAC AATTGCTGTT 120  
AGAGACTCTG CTTCAAATGG GAGCATTTC ATCACTTACG GACTTTTTCG TGGGGAGAGT 180  
AGTGAAGAAT TGAGTCACGG ACTTGCAGAA CCAAAGAAAA AGTTTGCAGT TTTAGAGATA 240  
5 CTGAATAATT CTTCCCAAAA AACTCTGCAT TCGGTGACTA TCCTGTTTCT GGTCTGAGT 300  
TTGATCACGT CGCTGCTGAG CTCTGGGTTT ACCTTCTACA ACAGCATCAG CAACCCTTAC 360  
CAGACATTCC TGGGGCCGAC GGGGGTGTAC ACCTGGAACG GGCTCGGTGC ATCCTTCGTT 420  
TTTGTGACCA TGATACTGTT TGTGGCGAAC ACGCAGTCCA ACCAACTCTC CGAAGAGTTG 480  
TTCCAAATGC TTTACCCGGC AACCACCAGT AAAGGAACGA CCCACAGTTA CGGATACTCG 540  
10 TTCTGGCTCA TACTGCTCGT CATTCTTCTA AATATAGTCA CTGTAACCAT CATCATTTTC 600  
TACCAGAAGG CCAGATACCA GCGGAAGCAG GAGCAGAGAA AGCCAATGGA ATATGCTCCA 660  
AGGGACGGAA TTTTATTC 678

## 15 (2) INFORMATION FOR SEQ ID NO: 34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 387  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
20 (D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
25 (B) CELL KIND: Liver  
(D) CLONE NAME: HP10432

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

30

ATGGCTCGGG GCTCGCTGCG CCGGTTGCTG CGGCTCCTCG TGCTGGGGCT CTGGCTGGCG 60  
TTGCTGCGCT CCGTGGCCGG GGAGCAAGCG CCAGGCACCG CCCCCTGCTC CCGCGGCAGC 120  
TCCTGGAGCG CGGACCTGGA CAAGTGCATG GACTGCGCGT CTTGCAGGGC GCGACCGCAC 180  
AGCGACTTCT GCCTGGGCTG CGCTGCAGCA CCTCCTGCCC CCTTCCGGCT GCTTTGGCCC 240  
35 ATCCTTGGGG GCGCTCTGAG CCTGACCTTC GTGCTGGGGC TGCTTTCTGG CTTTTTGGTC 300  
TGGAGACGAT GCCGCAGGAG AGAGAAGTTC ACCACCCCCA TAGAGGAGAC CGGCGGAGAG 360  
GGCTGCCCAG CTGTGGCGCT GATCCAG 387

## (2) INFORMATION FOR SEQ ID NO: 35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 489  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Liver  
(D) CLONE NAME: HP10433

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

ATGCGACGGC TGCTGATCCC TCTGGCCCTG TGGCTGGGCG CGGTGGGCGT GGGCGTCGCC 60  
GAGCTCACGG AAGCCCAGCG CCGGGGCCTG CAGGTGGCCC TGGAGGAATT TCACAAGCAC 120  
CCGCCCCGTGC AGTGGGCCTT CCAGGAGACC AGTGTGGAGA GCGCCGTGGA CACGCCCTTC 180  
CCAGCTGGAA TATTTGTGAG GCTGGAATTT AAGCTGCAGC AGACAAGCTG CCGGAAGAGG 240  
GACTGGAAGA AACCCGAGTG CAAAGTCAGG CCCAATGGGA GGAAACGGAA ATGCCTGGCC 300  
TGCATCAAAC TGGGCTCTGA GGACAAAGTT CTGGGCCGGT TGGTCCACTG CCCCATAGAG 360  
ACCCAAGTTC TCGGGAGGC TGAGGAGCAC CAGGAGACCC AGTGCCTCAG GGTGCAGCGG 420  
GCTGGTGAGG ACCCCCACAG CTTCTACTTC CCTGGACAGT TCGCCTTCTC CAAGGCCCTG 480  
CCCCGCAGC 489

## (2) INFORMATION FOR SEQ ID NO: 36:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 579  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP10480

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

ATGATCCGCT GCGGCCTGGC CTGCGAGCGC TGCCGCTGGA TCCTGCCCCT GCTCCTACTC 60  
 AGCGCCATCG CCTTCGACAT CATCGCGCTG GCCGGCCGCG GCTGGTTGCA GTCTAGCGAC 120  
 5 CACGGCCAGA CGTCCTCGCT GTGGTGGAAA TGCTCCCAAG AGGGCGGCGG CAGCGGGTCC 180  
 TACGAGGAGG GCTGTCAGAG CCTCATGGAG TACGCGTGGG GTAGAGCAGC GGCTGCCATG 240  
 CTCTTCTGTG GCTTCATCAT CCTGGTGATC TGTTTCATCC TCTCCTTCTT CGCCCTCTGT 300  
 GGACCCCAAGA TGCTTGTCTT CCTGAGAGTG ATTGGAGGTC TCCTTGCCTT GGCTGCTGTG 360  
 TTCCAGATCA TCTCCCTGGT AATTTACCCC GTGAAGTACA CCCAGACCTT CACCCTTCAT 420  
 10 GCCAACCGTG CTGTCACCTA CATCTATAAC TGGGCCTACG GCTTTGGGTG GGCAGCCACG 480  
 ATTATCCTGA TCGGCTGTGC CTTCTTCTTC TGCTGCCTCC CCAACTACGA AGATGACCTT 540  
 CTGGGCAATG CCAAGCCCAG GTACTTCTAC ACATCTGCC 579

15 (2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1502  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 20 (D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 25 (B) CELL KIND: Liver  
 (D) CLONE NAME: HP01263

(ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 30 (B) EXISTENCE POSITION: 37.. 1185  
 (C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

35 ACAAACCTGAC CCATCCTGGG CTTGTTCTC CACAGA ATG GGT CTG CTC CTT CCC 54  
 Met Gly Leu Leu Leu Pro  
 1 5  
 CTG GCA CTC TGC ATC CTA GTC CTG TGC TGC GGA GCA ATG TCT CCA CCC 102

0944553-120100  
 00T02T-3525450



120

Leu Ala Leu Cys Ile Leu Val Leu Cys Cys Gly Ala Met Ser Pro Pro  
 10 15 20  
 CAG CTG GCC CTC AAC CCC TCG GCT CTG CTC TCC CGG GGC TGC AAT GAC 150  
 Gln Leu Ala Leu Asn Pro Ser Ala Leu Leu Ser Arg Gly Cys Asn Asp  
 5 25 30 35  
 TCC GAT GTG CTG GCA GTT GCA GGC TTT GCC CTG CGG GAT ATT AAC AAA 198  
 Ser Asp Val Leu Ala Val Ala Gly Phe Ala Leu Arg Asp Ile Asn Lys  
 40 45 50  
 GAC AGA AAG GAT GGC TAT GTG CTG AGA CTC AAC CGA GTG AAC GAC GCC 246  
 10 Asp Arg Lys Asp Gly Tyr Val Leu Arg Leu Asn Arg Val Asn Asp Ala  
 55 60 65 70  
 CAG GAA TAC AGA CGG GGT GGC CTG GGA TCT CTG TTC TAT CTT ACA CTG 294  
 Gln Glu Tyr Arg Arg Gly Gly Leu Gly Ser Leu Phe Tyr Leu Thr Leu  
 75 80 85  
 15 GAT GTG CTA GAG ACT GAC TGC CAT GTG CTC AGA AAG AAG GCA TGG CAA 342  
 Asp Val Leu Glu Thr Asp Cys His Val Leu Arg Lys Lys Ala Trp Gln  
 90 95 100  
 GAC TGT GGA ATG AGG ATA TTT TTT GAA TCA GTT TAT GGT CAA TGC AAA 390  
 Asp Cys Gly Met Arg Ile Phe Phe Glu Ser Val Tyr Gly Gln Cys Lys  
 20 105 110 115  
 GCA ATA TTT TAT ATG AAC AAC CCA AGT AGA GTT CTC TAT TTA GCT GCT 438  
 Ala Ile Phe Tyr Met Asn Asn Pro Ser Arg Val Leu Tyr Leu Ala Ala  
 120 125 130  
 TAT AAC TGT ACT CTT CGC CCA GTT TCA AAA AAA AAG ATT TAC ATG ACG 486  
 25 Tyr Asn Cys Thr Leu Arg Pro Val Ser Lys Lys Lys Ile Tyr Met Thr  
 135 140 145 150  
 TGC CCT GAC TGC CCA AGC TCC ATA CCC ACT GAC TCT TCC AAT CAC CAA 534  
 Cys Pro Asp Cys Pro Ser Ser Ile Pro Thr Asp Ser Ser Asn His Gln  
 155 160 165  
 30 GTG CTG GAG GCT GCC ACC GAG TCT CTT GCG AAA TAC AAC AAT GAG AAC 582  
 Val Leu Glu Ala Ala Thr Glu Ser Leu Ala Lys Tyr Asn Asn Glu Asn  
 170 175 180  
 ACA TCC AAG CAG TAT TCT CTC TTC AAA GTC ACC AGG GCT TCT AGC CAG 630  
 Thr Ser Lys Gln Tyr Ser Leu Phe Lys Val Thr Arg Ala Ser Ser Gln  
 35 185 190 195  
 TGG GTG GTC GGC CCT TCT TAC TTT GTG GAA TAC TTA ATT AAA GAA TCA 678  
 Trp Val Val Gly Pro Ser Tyr Phe Val Glu Tyr Leu Ile Lys Glu Ser  
 200 205 210

69445258 120169

121

	CCA	TGT	ACT	AAA	TCC	CAG	GCC	AGC	AGC	TGT	TCA	CTT	CAG	TCC	TCC	GAC	726
	Pro	Cys	Thr	Lys	Ser	Gln	Ala	Ser	Ser	Cys	Ser	Leu	Gln	Ser	Ser	Asp	
	215				220					225					230		
	TCT	GTG	CCT	GTT	GGT	CTT	TGC	AAA	GGT	TCT	CTG	ACT	CGA	ACA	CAC	TGG	774
5	Ser	Val	Pro	Val	Gly	Leu	Cys	Lys	Gly	Ser	Leu	Thr	Arg	Thr	His	Trp	
					235					240					245		
	GAA	AAG	TTT	GTC	TCT	GTG	ACT	TGT	GAC	TTC	TTT	GAA	TCA	CAG	GCT	CCA	822
	Glu	Lys	Phe	Val	Ser	Val	Thr	Cys	Asp	Phe	Phe	Glu	Ser	Gln	Ala	Pro	
				250					255					260			
10	GCC	ACT	GGA	AGT	GAA	AAC	TCT	GCT	GTT	AAC	CAG	AAA	CCT	ACA	AAC	CTT	870
	Ala	Thr	Gly	Ser	Glu	Asn	Ser	Ala	Val	Asn	Gln	Lys	Pro	Thr	Asn	Leu	
				265					270					275			
	CCC	AAG	GTG	GAA	GAA	TCC	CAG	CAG	AAA	AAC	ACC	CCC	CCA	ACA	GAC	TCC	918
	Pro	Lys	Val	Glu	Glu	Ser	Gln	Gln	Lys	Asn	Thr	Pro	Pro	Thr	Asp	Ser	
15		280				285					290						
	CCC	TCC	AAA	GCT	GGG	CCA	AGA	GGA	TCT	GTC	CAA	TAT	CTT	CCT	GAC	TTG	966
	Pro	Ser	Lys	Ala	Gly	Pro	Arg	Gly	Ser	Val	Gln	Tyr	Leu	Pro	Asp	Leu	
	295				300					305				310			
	GAT	GAT	AAA	AAT	TCC	CAG	GAA	AAG	GGC	CCT	CAG	GAG	GCC	TTT	CCT	GTG	1014
20	Asp	Asp	Lys	Asn	Ser	Gln	Glu	Lys	Gly	Pro	Gln	Glu	Ala	Phe	Pro	Val	
				315						320				325			
	CAT	CTG	GAC	CTA	ACC	ACG	AAT	CCC	CAG	GGA	GAA	ACC	CTG	GAT	ATT	TCC	1062
	His	Leu	Asp	Leu	Thr	Thr	Asn	Pro	Gln	Gly	Glu	Thr	Leu	Asp	Ile	Ser	
				330					335					340			
25	TTC	CTC	TTC	CTG	GAG	CCT	ATG	GAG	GAG	AAG	CTG	GTT	GTC	CTG	CCT	TTC	1110
	Phe	Leu	Phe	Leu	Glu	Pro	Met	Glu	Glu	Lys	Leu	Val	Val	Leu	Pro	Phe	
				345					350					355			
	CCC	AAA	GAA	AAA	GCA	CGC	ACT	GCT	GAG	TGC	CCA	GGG	CCA	GCC	CAG	AAT	1158
	Pro	Lys	Glu	Lys	Ala	Arg	Thr	Ala	Glu	Cys	Pro	Gly	Pro	Ala	Gln	Asn	
30		360				365					370						
	GCC	AGC	CCT	CTT	GTC	CTT	CCG	CCA	TGAGAATCAC	ACAGAGTCTT	CTGTAGGG						1210
	Ala	Ser	Pro	Leu	Val	Leu	Pro	Pro									
	375					380											
	GTATGGTGCG	CCGCATGACA	TGGGAGGCCA	TGGGGACGAT	GGACAGAGAC	AGAGCGTGCA											1270
35	CACGTAGAGT	GGCTAGTGAA	GGACGCCTTT	TTGACTCTTC	TTGGTCTCAG	CATGTTGACT											1330
	GGGATTGGAA	ATAATGAGAC	TGAGCCCTCG	GCTTGGGCTG	CACTCTACCC	TGTACACTGC											1390
	CTTGTACCCT	GAGCTGCATC	ACCTCCTAAA	CTGAGCAGTC	TCATACCATG	GAGAGATGCC											1450
	TCTCTTATGT	CTTCAGCCAC	TCACTTATAA	AGATACTTAT	CTTTTCAGCA	GT											1502

09445258 120199

## (2) INFORMATION FOR SEQ ID NO: 38:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1349

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Liver

(D) CLONE NAME: HP01299

## (ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 111.. 1064

(C) CHARACTERIZATION METHOD: E

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```

AGCAGTTGGG GCAGGAGGAA GCCGACTGCT GCCTGGTCTG CAAAGAAGTC CTTTCAAGTC      60
TCTAGGACTG GACTCTTCCT AAGCAAGTCC GAGAAGGAAG CACCCTCACT ATG TGG      116
                                         Met Trp
                                         1
CTC TAC CTG GCG GCC TTC GTG GGC CTG TAC TAC CTT CTG CAC TGG TAC
164
Leu Tyr Leu Ala Ala Phe Val Gly Leu Tyr Tyr Leu Leu His Trp Tyr
30      5              10              15
CGG GAG AGG CAG GTG GTG AGC CAC CTC CAA GAC AAG TAT GTC TTT ATC      212
Arg Glu Arg Gln Val Val Ser His Leu Gln Asp Lys Tyr Val Phe Ile
      20              25              30
ACG GGC TGT GAC TCG GGC TTT GGG AAC CTG CTG GCC AGA CAG CTG GAT      260
35 Thr Gly Cys Asp Ser Gly Phe Gly Asn Leu Leu Ala Arg Gln Leu Asp
      35              40              45              50
GCA CGA GGC TTG AGA GTG CTG GCT GCG TGT CTG ACG GAG AAG GGG GCC      308
Ala Arg Gly Leu Arg Val Leu Ala Ala Cys Leu Thr Glu Lys Gly Ala

```

09445258 120199

123

				55					60					65			
				GAG CAG CTG AGG GGC CAG ACG TCT GAC AGG CTG GAG ACG GTG ACC CTG													356
				Glu Gln Leu Arg Gly Gln Thr Ser Asp Arg Leu Glu Thr Val Thr Leu													
				70					75					80			
5				GAT GTT ACC AAG ATG GAG AGC ATC GCT GCA GCT ACT CAG TGG GTG AAG													404
				Asp Val Thr Lys Met Glu Ser Ile Ala Ala Ala Thr Gln Trp Val Lys													
				85					90					95			
				GAG CAT GTG GGG GAC AGA GGA CTC TGG GGA CTG GTG AAC AAT GCA GGC													452
				Glu His Val Gly Asp Arg Gly Leu Trp Gly Leu Val Asn Asn Ala Gly													
10				100					105					110			
				ATT CTT ACA CCA ATT ACC TTA TGT GAG TGG CTG AAC ACT GAG GAC TCT													500
				Ile Leu Thr Pro Ile Thr Leu Cys Glu Trp Leu Asn Thr Glu Asp Ser													
				115					120					125			130
				ATG AAT ATG CTC AAA GTG AAC CTC ATT GGT GTG ATC CAG GTG ACC TTG													548
15				Met Asn Met Leu Lys Val Asn Leu Ile Gly Val Ile Gln Val Thr Leu													
				135					140					145			
				AGC ATG CTT CCT TTG GTG AGG AGA GCA CGG GGA AGA ATT GTC AAT GTC													596
				Ser Met Leu Pro Leu Val Arg Arg Ala Arg Gly Arg Ile Val Asn Val													
				150					155					160			
20				TCC AGC ATT CTG GGA AGA GTT GCT TTC TTT GTA GGA GGC TAC TGT GTC													644
				Ser Ser Ile Leu Gly Arg Val Ala Phe Phe Val Gly Gly Tyr Cys Val													
				165					170					175			
				TCC AAG TAT GGA GTG GAA GCC TTT TCA GAT ATT CTG AGG CGT GAG ATT													692
				Ser Lys Tyr Gly Val Glu Ala Phe Ser Asp Ile Leu Arg Arg Glu Ile													
25				180					185					190			
				CAA CAT TTT GGG GTG AAA ATC AGC ATA GTT GAA CCT GGC TAC TTC AGA													740
				Gln His Phe Gly Val Lys Ile Ser Ile Val Glu Pro Gly Tyr Phe Arg													
				195					200					205			210
				ACG GGA ATG ACA AAC ATG ACA CAG TCC TTA GAG CGA ATG AAG CAA AGT													788
30				Thr Gly Met Thr Asn Met Thr Gln Ser Leu Glu Arg Met Lys Gln Ser													
				215					220					225			
				TGG AAA GAA GCC CCC AAG CAT ATT AAG GAG ACC TAT GGA CAG CAG TAT													836
				Trp Lys Glu Ala Pro Lys His Ile Lys Glu Thr Tyr Gly Gln Gln Tyr													
				230					235					240			
35				TTT GAT GCC CTT TAC AAT ATC ATG AAG GAA GGG CTG TTG AAT TGT AGC													884
				Phe Asp Ala Leu Tyr Asn Ile Met Lys Glu Gly Leu Leu Asn Cys Ser													
				245					250					255			
				ACA AAC CTG AAC CTG GTC ACT GAC TGC ATG GAA CAT GCT CTG ACA TCG													932

09445258-120199

124

Thr Asn Leu Asn Leu Val Thr Asp Cys Met Glu His Ala Leu Thr Ser  
 260 265 270  
 GTG CAT CCG CGA ACT CGA TAT TCA GCT GGC TGG GAT GCT AAA TTT TTC 980  
 Val His Pro Arg Thr Arg Tyr Ser Ala Gly Trp Asp Ala Lys Phe Phe  
 5 275 280 285 290  
 TTC ATC CCT CTA TCT TAT TTA CCT ACA TCA CTG GCA GAC TAC ATT TTG 1028  
 Phe Ile Pro Leu Ser Tyr Leu Pro Thr Ser Leu Ala Asp Tyr Ile Leu  
 295 300 305  
 ACT AGA TCT TGG CCC AAA CCA GCC CAG GCA GTC TAAAGAAAAC TGGGTTGGT 1080  
 10 Thr Arg Ser Trp Pro Lys Pro Ala Gln Ala Val  
 310 315  
 GCTTCTTGGA ATGAAGGCAA AAATCTGAAA TTGTTAGTGT CTCAGTAATC CTGATTTAGA 1140  
 ACCCAGGCTT TTTGTAACAA TGTGTTTTCT TGCCTAAATT CATTTATCTG GCATCATCAG 1200  
 AGTACTAACA TGTTTATATT TCAGATATCC AAAGCTTACC ACTTTAGGTG ATGAATCTTT 1260  
 15 ACTATTTTAG CCCTTTTTTG ATGAGACTAT TTGTCTAAAG TGAATCATTT GTTCTTGCCT 1320  
 TATTAAACAG AGTAGATGGA AAACAATTT 1349

## (2) INFORMATION FOR SEQ ID NO: 39:

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1643
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 25 (ii) SEQUENCE KIND: cDNA to mRNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Homo sapiens*
  - (B) CELL KIND: Liver
  - 30 (D) CLONE NAME: HP01347
- (ix) SEQUENCE CHARACTERISTICS:
- (A) CHARACTERIZATION CODE: CDS
  - (B) EXISTENCE POSITION: 25.. 915
  - 35 (C) CHARACTERIZATION METHOD: E
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

094455508-120199  
 66T02T 85254460

125

AACATCTGGG GACAGCGGGA AAAC ATG AGT GAC TCC AAG GAA CCA AGG GTG 51  
Met Ser Asp Ser Lys Glu Pro Arg Val  
1 5  
CAG CAG CTG GGC CTC CTG GGG TGT CTT GGC CAT GGC GCC CTG GTG CTG 99  
5 Gln Gln Leu Gly Leu Leu Gly Cys Leu Gly His Gly Ala Leu Val Leu  
10 15 20 25  
CAA CTC CTC TCC TTC ATG CTC TTG GCT GGG GTC CTG GTG GCC ATC CTT 147  
Gln Leu Leu Ser Phe Met Leu Leu Ala Gly Val Leu Val Ala Ile Leu  
30 35 40  
10 GTC CAA GTG TCC AAG GTC CCC AGC TCC CTA AGT CAG GAA CAA TCC GAG 195  
Val Gln Val Ser Lys Val Pro Ser Ser Leu Ser Gln Glu Gln Ser Glu  
45 50 55  
CAA GAC GCA ATC TAC CAG AAC CTG ACC CAG CTT AAA GCT GCA GTG GGT 243  
Gln Asp Ala Ile Tyr Gln Asn Leu Thr Gln Leu Lys Ala Ala Val Gly  
15 60 65 70  
GAG CTC TCA GAG AAA TCC AAG CTG CAG GAG ATC TAC CAG GAG CTG ACC 291  
Glu Leu Ser Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr  
75 80 85  
CAG CTG AAG GCT GCA GTG GGT GAG TTG CCA GAG AAA TCC AAG CTG CAG 339  
20 Gln Leu Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln  
90 95 100 105  
GAG ATC TAC CAG GAG CTG ACC CGG CTG AAG GCT GCA GTG GGT GAG TTG 387  
Glu Ile Tyr Gln Glu Leu Thr Arg Leu Lys Ala-Ala Val Gly Glu Leu  
110 115 120  
25 CCA GAG AAA TCC AAG CTG CAG GAG ATC TAC CAG GAG CTG ACC CGG CTG 435  
Pro Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr Arg Leu  
125 130 135  
AAG GCT GCA GTG GGT GAG TTG CCA GAG AAA TCC AAG CTG CAG GAG ATC 483  
Lys Ala Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln Glu Ile  
30 140 145 150  
TAC CAG GAG CTG ACC CGG CTG AAG GCT GCA GTG GGT GAG TTG CCA GAG 531  
Tyr Gln Glu Leu Thr Arg Leu Lys Ala Ala Val Gly Glu Leu Pro Glu  
155 160 165  
AAA TCC AAG CTG CAG GAG ATC TAC CAG GAG CTG ACG GAG CTG AAG GCT 579  
35 Lys Ser Lys Leu Gln Glu Ile Tyr Gln Glu Leu Thr Glu Leu Lys Ala  
170 175 180 185  
GCA GTG GGT GAG TTG CCA GAG AAA TCC AAG CTG CAG GAG ATC TAC CAG 627  
Ala Val Gly Glu Leu Pro Glu Lys Ser Lys Leu Gln Glu Ile Tyr Gln

09445258 120109

126

		190		195		200	
		GAG CTG ACC CAG CTG AAG GCT GCA GTG GGT GAG TTG CCA GAC CAG TCC					675
		Glu Leu Thr Gln Leu Lys Ala Ala Val Gly Glu Leu Pro Asp Gln Ser					
		205		210		215	
5		AAG CAG CAG CAA ATC TAT CAA GAA CTG ACC GAT TTG AAG ACT GCA TTT					723
		Lys Gln Gln Gln Ile Tyr Gln Glu Leu Thr Asp Leu Lys Thr Ala Phe					
		220		225		230	
		GAA CGC CTG TGC CGC CAC TGT CCC AAG GAC TGG ACA TTC TTC CAA GGA					771
		Glu Arg Leu Cys Arg His Cys Pro Lys Asp Trp Thr Phe Phe Gln Gly					
10		235		240		245	
		AAC TGT TAC TTC ATG TCT AAC TCC CAG CGG AAC TGG CAC GAC TCC GTC					819
		Asn Cys Tyr Phe Met Ser Asn Ser Gln Arg Asn Trp His Asp Ser Val					
		250		255		260	
		ACC GCC TGC CAG GAA GTG AGG GCC CAG CTC GTC GTA ATC AAA ACT GCT					867
15		Thr Ala Cys Gln Glu Val Arg Ala Gln Leu Val Val Ile Lys Thr Ala					
		270		275		280	
		GAG GAG CAG CTT CCA GCG GTA CTG GAA CAG TGG AGA ACC CAA CAA					912
		Glu Glu Gln Leu Pro Ala Val Leu Glu Gln Trp Arg Thr Gln Gln					
		285		290		295	
20		TAGCGGGAAT GAAGACTGTG CGGAATTTAG TGGCAGTGGC TGGAACGACA ATCGATGT					970
		GACGTTGACA ATTACTGGAT CTGCAAAAAG CCCGCAGCCT GCTTCAGAGA CGAATAGTTG					1030
		TTTCCTTGCT AGCCTCAGCC TCCATTGTGG TATAGCAGAA CTTACCCAC TTGTAAGCCA					1090
		GCGCTTCTTC TCTCCATCCT TGGACCTTCA CAAATGCCCT GAGACGGTTC TCTGTTTCGAT					1150
		TTTTCATCCC CTATGAACCT GGGTCTTATT CTGTCCTTCT GATGCCTCCA AGTTTCCCTG					1210
25		GTGTAGAGCT TGTGTTCTTG GCCCATCCTT GGAGCTTTAT AAGTGACCTG AGTGGGATGC					1270
		ATTTAGGGGG CGGGCTTGGT ATGTTGTATG AATCCACTCT CTGTTCTTTT TGGAGATTAG					1330
		ACTATTTGGA TTCATGTGTA GCTGCCCTGT CCCCTGGGGC TTTATCTCAT CCATGCAAAC					1390
		TACCATCTGC TCAACTTCCA GCTACACCCC GTGCACCCTT TTGACTGGGG ACTTGCTGGT					1450
		TGAAGGAGCT CATCTTGCAG GCTGGAAGCA CCAGGGAATT AATTCCCCCA GTCAACCAAT					1510
30		GGCATCCAGA GAGGGCATGG AGGCTCCATA CAACCTCTTC CACCCCCACA TCTTTCTTTG					1570
		TCCTATACAT GTCTTCCATT TGGCTGTTTC TGAGTTGTAG CCTTTATAAT AAAGTGGTAA					1630
		ATGTTGTAAC TGC					1643

35 (2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 729

(B) TYPE: Nucleic acid

667027 852460

127

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP01440

10 (ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 38.. 631

(C) CHARACTERIZATION METHOD: E

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

```

ACTTTCCTC ACCGCCTGTC CTTCTGACA CCTCACC ATG TGT ACG GGA AAA TGT      55
                                Met Cys Thr Gly Lys Cys
                                1           5
20 GCC CGC TGT GTG GGG CTC TCC CTC ATT ACC CTC TGC CTC GTC TGC ATT      103
   Ala Arg Cys Val Gly Leu Ser Leu Ile Thr Leu Cys Leu Val Cys Ile
           10           15           20
   GTG GCC AAC GCC CTC CTG CTG GTA CCT AAT GGG GAG ACC TCC TGG ACC      151
   Val Ala Asn Ala Leu Leu Leu Val Pro Asn Gly Glu Thr Ser Trp Thr
25           25           30           35
   AAC ACC AAC CAT CTC AGC TTG CAA GTC TGG CTC ATG GGC GGC TTC ATT      199
   Asn Thr Asn His Leu Ser Leu Gln Val Trp Leu Met Gly Gly Phe Ile
           40           45           50
   GGC GGC GGC CTA ATG GTA CTG TGT CCG GGC ATT GCA GCC GTT CGG GCA      247
30 Gly Gly Gly Leu Met Val Leu Cys Pro Gly Ile Ala Ala Val Arg Ala
           55           60           65           70
   GGC GGC AAG GGC TGC TGT GGT GCT GGC TGC TGT GGA AAC CGC TGC AGG      295
   Gly Gly Lys Gly Cys Cys Gly Ala Gly Cys Cys Gly Asn Arg Cys Arg
           75           80           85
35 ATG CTG CGC TCG GTC TTC TCC TCG GCG TTC GGG GTG CTT GGT GCC ATC      343
   Met Leu Arg Ser Val Phe Ser Ser Ala Phe Gly Val Leu Gly Ala Ile
           90           95           100
   TAC TGC CTC TCG GTG TCT GGA GCT GGC CTC CGA AAT GGA CCC AGA TGC      391

```

09445258-120169



128

Tyr Cys Leu Ser Val Ser Gly Ala Gly Leu Arg Asn Gly Pro Arg Cys  
 105 110 115  
 TTA ATG AAC GGC GAG TGG GGC TAC CAC TTC GAA GAC ACC GCG GGA GCT 439  
 Leu Met Asn Gly Glu Trp Gly Tyr His Phe Glu Asp Thr Ala Gly Ala  
 5 120 125 130  
 TAC TTG CTC AAC CGC ACT CTA TGG GAT CGG TGC GAG GCG CCC CCT CGC 487  
 Tyr Leu Leu Asn Arg Thr Leu Trp Asp Arg Cys Glu Ala Pro Pro Arg  
 135 140 145 150  
 GTG GTC CCC TGG AAT GTG ACG CTC TTC TCG CTG CTG GTG GCC GCC TCC 535  
 10 Val Val Pro Trp Asn Val Thr Leu Phe Ser Leu Leu Val Ala Ala Ser  
 155 160 165  
 TGC CTG GAG ATA GTA CTG TGT GGG ATC CAG CTG GTG AAC GCG ACC ATT 583  
 Cys Leu Glu Ile Val Leu Cys Gly Ile Gln Leu Val Asn Ala Thr Ile  
 170 175 180  
 15 GGT GTC TTC TGC GGC GAT TGC AGG AAA AAA CAG GAC ACC CCT CAC TG 630  
 Gly Val Phe Cys Gly Asp Cys Arg Lys Lys Gln Asp Thr Pro His  
 185 190 195  
 AGGCTCCACT GACCGCCGGG TTACACCTGC TCCTTCCTGG ACGCCTACCT GGCTCGCTCA 690  
 CTCCCTTGCT CGCTAGAATA AACTGCTTTG CGCTCTCTT 729

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1322

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP01526

(ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 84.. 749

(C) CHARACTERIZATION METHOD: E

551021 825460

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

GAGCCGCAGG TCTGGGCTGC AGTAGGTCCC GGCAACCGCA GGCTCGCGGC GGGCGCTGGG 60  
 CGCGGGATCC GACTCTAGTC GTA ATG GAG GCG GGC GGC TTT CTG GAC TCG CTC 113  
 5 Met Glu Ala Gly Gly Phe Leu Asp Ser Leu  
 1 5 10  
 ATT TAC GGA GCA TGC GTG GTC TTC ACC CTT GGC ATG TTC TCC GCC GGC 161  
 Ile Tyr Gly Ala Cys Val Val Phe Thr Leu Gly Met Phe Ser Ala Gly  
 15 20 25  
 10 CTC TCG GAC CTC AGG CAC ATG CGA ATG ACC CGG AGT GTG GAC AAC GTC 209  
 Leu Ser Asp Leu Arg His Met Arg Met Thr Arg Ser Val Asp Asn Val  
 30 35 40  
 CAG TTC CTG CCC TTT CTC ACC ACG GAA GTC AAC AAC CTG GGC TGG CTG 257  
 Gln Phe Leu Pro Phe Leu Thr Thr Glu Val Asn Asn Leu Gly Trp Leu  
 15 45 50 55  
 AGT TAT GGG GCT TTG AAG GGA GAC GGG ATC CTC ATC GTC GTC AAC ACA 305  
 Ser Tyr Gly Ala Leu Lys Gly Asp Gly Ile Leu Ile Val Val Asn Thr  
 60 65 70  
 GTG GGT GCT GCG CTT CAG ACC CTG TAT ATC TTG GCA TAT CTG CAT TAC 353  
 Val Gly Ala Ala Leu Gln Thr Leu Tyr Ile Leu Ala Tyr Leu His Tyr  
 20 75 80 85 90  
 TGC CCT CGG AAG CGT GTT GTG CTC CTA CAG ACT GCA ACC CTG CTA GGG 401  
 Cys Pro Arg Lys Arg Val Val Leu Leu Gln Thr Ala Thr Leu Leu Gly  
 95 100 105  
 25 GTC CTT CTC CTG GGT TAT GGC TAC TTT TGG CTC CTG GTA CCC AAC CCT 449  
 Val Leu Leu Leu Gly Tyr Gly Tyr Phe Trp Leu Leu Val Pro Asn Pro  
 110 115 120  
 GAG GCC CGG CTT CAG CAG TTG GGC CTC TTC TGC AGT GTC TTC ACC ATC 497  
 Glu Ala Arg Leu Gln Gln Leu Gly Leu Phe Cys Ser Val Phe Thr Ile  
 30 125 130 135  
 AGC ATG TAC CTC TCA CCA CTG GCT GAC TTG GCT AAG GTG ATT CAA ACT 545  
 Ser Met Tyr Leu Ser Pro Leu Ala Asp Leu Ala Lys Val Ile Gln Thr  
 140 145 150  
 AAA TCA ACC CAA TGT CTC TCC TAC CCA CTC ACC ATT GCT ACC CTT CTC 593  
 35 Lys Ser Thr Gln Cys Leu Ser Tyr Pro Leu Thr Ile Ala Thr Leu Leu  
 155 160 165 170  
 ACC TCT GCC TCC TGG TGC CTC TAT GGG TTT CGA CTC AGA GAT CCC TAT 641  
 Thr Ser Ala Ser Trp Cys Leu Tyr Gly Phe Arg Leu Arg Asp Pro Tyr

09445258.120199

130

	175	180	185	
	ATC ATG GTG TCC AAC TTT CCA GGA ATC GTC ACC AGC TTT ATC CGC TTC			689
	Ile Met Val Ser Asn Phe Pro Gly Ile Val Thr Ser Phe Ile Arg Phe			
	190	195	200	
5	TGG CTT TTC TGG AAG TAC CCC CAG GAG CAA GAC AGG AAC TAC TGG CTC			737
	Trp Leu Phe Trp Lys Tyr Pro Gln Glu Gln Asp Arg Asn Tyr Trp Leu			
	205	210	215	
	CTG CAA ACC TGAGGCTGCT CATCTGACCA CTGGGCACCT TAGTGCCAAC CTGA			790
	Leu Gln Thr			
10	220			
	ACCAAAGAGA CCTCCTTGTT TCAGCTGGGC CTGCTGTCCA GCTTCCCAGG TGCAGTGGGT			850
	TGTGGGAACA AGAGATGACT TTGAGGATAA AAGGACCAAA GAAAAAGCTT TACTTAGATG			910
	ATTGATTGGG GCCTAGGAGA TGAAATCACT TTTTATTTT TAGAGATTTT TTTTTTTAAT			970
	TTTGGAGGTT GGGGTGCAAT CTTTAGAATA TGCCTTAAAA GGCCGGGCGC GGTGGCTCAC			1030
15	GCCTGTAATC CCAGCACTTT GGGAGGCCAA GGTGGGCGGA TCGCCTGAGG TCAGGAGTTC			1090
	AAGACCAACC TGACTAACAT GGTGAAACCC CATCTCTACT AAAAATACAA AATTAGCCAG			1150
	GCATGATGGC ACATGCCTGT AATCCCAGAT ACTTGGGAGG CTGAGGCAGG AGAATTGCTT			1210
	GAACCCAGGA GGTGGAGGTT GCAGTGAGCT GAGATCGTGC CATTGTGATA TGAATATGCC			1270
	TTATATGCTG ATATGAATAT GCCTTAAAT AAAGTGTTC CCACCCCTGC CC			1322

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3045

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10230

(ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 191.. 946

(C) CHARACTERIZATION METHOD: E

131

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

GTTTCGCCTC AGAAGGCTGC CTCGCTGGTC CGAATTCGGT GGCGCCACGT CCGCCCGTCT 60  
 CCGCCTTCTG CATCGCGGCT TCGGCGGCTT CCACCTAGAC ACCTAACAGT CGCGGAGCCG 120  
 5 GCCGCGTCGT GAGGGGGTCG GCACGGGGAG TCGGGCGGTC TTGTGCATCT TGGCTACCTG 180  
 TGGGTCGAAG ATG TCG GAC ATC GGA GAC TGG TTC AGG AGC ATC CCG GCG 229  
 Met Ser Asp Ile Gly Asp Trp Phe Arg Ser Ile Pro Ala  
 1 5 10  
 ATC ACG CGC TAT TGG TTC GCC GCC ACC GTC GCC GTG CCC TTG GTC GGC 277  
 10 Ile Thr Arg Tyr Trp Phe Ala Ala Thr Val Ala Val Pro Leu Val Gly  
 15 20 25  
 AAA CTC GGC CTC ATC AGC CCG GCC TAC CTC TTC CTC TGG CCC GAA GCC 325  
 Lys Leu Gly Leu Ile Ser Pro Ala Tyr Leu Phe Leu Trp Pro Glu Ala  
 30 35 40 45  
 15 TTC CTT TAT CGC TTT CAG ATT TGG AGG CCA ATC ACT GCC ACC TTT TAT 373  
 Phe Leu Tyr Arg Phe Gln Ile Trp Arg Pro Ile Thr Ala Thr Phe Tyr  
 50 55 60  
 TTC CCT GTG GGT CCA GGA ACT GGA TTT CTT TAT TTG GTC AAT TTA TAT 421  
 Phe Pro Val Gly Pro Gly Thr Gly Phe Leu Tyr Leu Val Asn Leu Tyr  
 20 65 70 75  
 TTC TTA TAT CAG TAT TCT ACG CGA CTT GAA ACA GGA GCT TTT GAT GGG 469  
 Phe Leu Tyr Gln Tyr Ser Thr Arg Leu Glu Thr Gly Ala Phe Asp Gly  
 80 85 90  
 AGG CCA GCA GAC TAT TTA TTC ATG CTC CTC TTT AAC TGG ATT TGC ATC 517  
 25 Arg Pro Ala Asp Tyr Leu Phe Met Leu Leu Phe Asn Trp Ile Cys Ile  
 95 100 105  
 GTG ATT ACT GGC TTA GCA ATG GAT ATG CAG TTG CTG ATG ATT CCT CTG 565  
 Val Ile Thr Gly Leu Ala Met Asp Met Gln Leu Leu Met Ile Pro Leu  
 110 115 120 125  
 30 ATC ATG TCA GTA CTT TAT GTC TGG GCC CAG CTG AAC AGA GAC ATG ATT 613  
 Ile Met Ser Val Leu Tyr Val Trp Ala Gln Leu Asn Arg Asp Met Ile  
 130 135 140  
 GTA TCA TTT TGG TTT GGA ACA CGA TTT AAG GCC TGC TAT TTA CCC TGG 661  
 Val Ser Phe Trp Phe Gly Thr Arg Phe Lys Ala Cys Tyr Leu Pro Trp  
 35 145 150 155  
 GTT ATC CTT GGA TTC AAC TAT ATC ATC GGA GGC TCG GTA ATC AAT GAG 709  
 Val Ile Leu Gly Phe Asn Tyr Ile Ile Gly Gly Ser Val Ile Asn Glu  
 160 165 170

0944558.120199

132

	CTT ATT GGA AAT CTG GTT GGA CAT CTT TAT TTT TTC CTA ATG TTC AGA	757
	Leu Ile Gly Asn Leu Val Gly His Leu Tyr Phe Phe Leu Met Phe Arg	
	175 180 185	
	TAC CCA ATG GAC TTG GGA GGA AGA AAT TTT CTA TCC ACA CCT CAG TTT	805
5	Tyr Pro Met Asp Leu Gly Gly Arg Asn Phe Leu Ser Thr Pro Gln Phe	
	190 195 200 205	
	TTG TAC CGC TGG CTG CCC AGT AGG AGA GGA GGA GTA TCA GGA TTT GGT	853
	Leu Tyr Arg Trp Leu Pro Ser Arg Arg Gly Gly Val Ser Gly Phe Gly	
	210 215 220	
10	GTG CCC CCT GCT AGC ATG AGG CGA GCT GCT GAT CAG AAT GGC GGA GGC	901
	Val Pro Pro Ala Ser Met Arg Arg Ala Ala Asp Gln Asn Gly Gly Gly	
	225 230 235	
	GGG AGA CAC AAC TGG GGC CAG GGC TTT CGA CTT GGA GAC CAG TGAAGGG	950
	Gly Arg His Asn Trp Gly Gln Gly Phe Arg Leu Gly Asp Gln	
15	240 245 250	
	GCGGCCTCGG GCAGCCGCTC CTCTCAAGCC ACATTTCCCTC CCAGTGCTGG GTGCGCTTAA	1010
	CAACTGCGTT CTGGCTAACA CTGTTGGACC TGACCCACAC TGAATGTAGT CTTTCAGTAC	1070
	GAGACAAAGT TTCTTAAATC CCGAAGAAAA ATATAAGTGT TCCACAAGTT TCACGATTCT	1130
	CATTCAAGTC CTTACTGCTG TGAAGAACAA ATACCAACTG TGCAAATTGC AAAACTGACT	1190
20	ACATTTTTTG GTGTCTTCTC TTCTCCCCTT TCCGTCTGAA TAATGGGTTT TAGCGGGTCC	1250
	TAGTCTGCTG GCATTGAGCT GGGGCTGGGT CACCAAACCC TTCCCAAAAG GACCCTTATC	1310
	TCTTTCTTGC ACACATGCCT CTCTCCCACT TTTCCCAACC CCCACATTTG CAACTAGAAG	1370
	AGGTTGCCCA TAAAAATTGCT CTGCCCTTGA CAGGTTCTGT TATTTATTGA CTTTGGCCAA	1430
	GGCTTGGTCA CAACAATCAT ATTCACGTAA TTTTCCCCCT TTGGTGGCAG AACTGTAGCA	1490
25	ATAGGGGGAG AAGACAAGCA GCGGATGAAG CGTTTTCTCA GCTTTTGGAA TTGCTTCGAC	1550
	CTGACATCCG TTGTAACCGT TTGCCACTTC TTCAGATATT TTTATAAAAA AGTACCACTG	1610
	AGTCAGTGAG GGCCACAGAT TGGTATTAAT GAGATACGAG GGTGTTGCT GGGTGTGTTGT	1670
	TTCTTGAGCT AAGTGATCAA GACTGTAGTG GAGTTGCAGC TAACATGGGT TAGGTTTAAA	1730
	CCGTGGGGGA TGCAACCCCT TTGCGTTTCA TATGTAGGCC TACTGGCTTT GTGTAGCTGG	1790
30	AGTAGTTGGG TTGCTTTGTG TTAGGAGGAT CCAGATCATG TTGGCTACAG GGAGATGCTC	1850
	TCTTTGAGAG GCTCCTGGGC ATTGATTCCA TTTCAATCTC ATTCTGGATA TGTGTTTATT	1910
	GAGTAAAGGA GGAGAGACCC TCATACGCTA TTAAATGTC ACTTTTTTGC CTATCCCCCG	1970
	TTTTTTGGTC ATGTTTCAAT TAATTGTGAG GAAGGCGCAG CTCCTCTCTG CACGTAGATC	2030
	ATTTTTTAAA GCTAATGTAA GCACATCTAA GGGAATAACA TGATTTAAGG TTGAAATGGC	2090
35	TTTAGAATCA TTTGGGTTTG AGGGTGTGTT ATTTTGAGTC ATGAATGTAC AAGCTCTGTG	2150
	AATCAGACCA GCTTAAATAC CCACACCTTT TTTTCGTAGG TGGGCTTTTC CTATCAGAGC	2210
	TTGGCTCATA ACCAAATAAA GTTTTTTGAA GGCCATGGCT TTTACACAG TTATTTTATT	2270
	TTATGACGTT ATCTGAAAGC AGACTGTTAG GAGCAGTATT GAGTGGCTGT CACACTTTGA	2330

044558 120199

133

GGCAACTAAA AAGGCTTCAA ACGTTTTGAT CAGTTTCTTT TCAGGAAACA TTGTGCTCTA 2390  
ACAGTATGAC TATTCTTTCC CCCACTCTTA AACAGTGTGA TGTGTGTTAT CCTAGGAAAT 2450  
GAGAGTTGGC AAACAACCTC TCATTTTGAA TAGAGTTTGT GTGTACCTCT CCATATTTAA 2510  
TTTATATGAT AAAATAGGTG GGGAGAGTCT GAACCTTAAC TGTCATGTTT TGTGTTCAT 2570  
5 CTGTGGCCAC AATAAAGTTT ACTTGTAATA TTTTAGAGGC CATTACTCCA ATTATGTTGC 2630  
ACGTACACTC ATTGTACAGG CGTGGAGACT CATTGTATGT ATAAGAATAT TCTGACAGTG 2690  
AGTGACCCGG AGTCTCTGGT GTACCCTCTT ACCAGTCAGC TGCCTGCGAG CAGTCATTTT 2750  
TTCCTAAAGG TTTACAAGTA TTTAGAACTC TTCAGTTCAG GGCAAAATGT TCATGAAGTT 2810  
ATTCTCTTA AACATGGTTA GGAAGCTGAT GACGTTATTG ATTTTGTCTG GATTATGTTT 2870  
10 CTGGAATAAT TTTACCAAAA CAAGCTATTT GAGTTTTGAC TTGACAAGGC AAAACATGAC 2930  
AGTGGATTCT CTTTACAAAT TGAAAAAAAA AATCCTTATT TTGTATAAAG GACTTCCCTT 2990  
TTTGTAAGT AATCCTTTTT ATTGGTAAAA ATTGTAAATT AAAATGTGCA ACTTG 3045

## 15 (2) INFORMATION FOR SEQ ID NO: 43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
20 (D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
25 (B) CELL KIND: Epidermoid carcinoma  
(C) CELL LINE: KB  
(D) CLONE NAME: HP10389

## (ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
30 (B) EXISTENCE POSITION: 63.. 383  
(C) CHARACTERIZATION METHOD: E

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

35

ATGACCTTCA CCGGGAGGCT GAGGTCGGAG TCCCGATTTT CTCCTGCTGC TGTGGCCCGG 60  
AC ATG GCG ACT CCC GGC CCT GTG ATT CCG GAG GTC CCC TTT GAA CCA 107  
Met Ala Thr Pro Gly Pro Val Ile Pro Glu Val Pro Phe Glu Pro

134

1 5 10 15  
TCG AAG CCT CCA GTC ATT GAG GGG CTG AGC CCC ACT GTT TAC AGG AAT 155  
Ser Lys Pro Pro Val Ile Glu Gly Leu Ser Pro Thr Val Tyr Arg Asn  
20 25 30  
5 CCA GAG AGT TTC AAG GAA AAG TTC GTT CGC AAG ACC CGC GAG AAC CCG 203  
Pro Glu Ser Phe Lys Glu Lys Phe Val Arg Lys Thr Arg Glu Asn Pro  
35 40 45  
GTG GTA CCC ATA GGT TGC CTG GCC ACG GCG GCC GCC CTC ACC TAC GGC 251  
Val Val Pro Ile Gly Cys Leu Ala Thr Ala Ala Ala Leu Thr Tyr Gly  
10 50 55 60  
CTC TAC TCC TTC CAC CGG GGC AAC AGC CAG CGC TCT CAG CTC ATG ATG 299  
Leu Tyr Ser Phe His Arg Gly Asn Ser Gln Arg Ser Gln Leu Met Met  
65 70 75  
CGC ACC CGG ATC GCC GCC CAG GGT TTC ACG GTC GCA GCC ATC TTG CTG 347  
15 Arg Thr Arg Ile Ala Ala Gln Gly Phe Thr Val Ala Ala Ile Leu Leu  
80 85 90 95  
GGT CTG GCT GTC ACT GCT ATG AAG TCT CGA CCC TAAGCCCAGG GTCTGGCCTT 400  
Gly Leu Ala Val Thr Ala Met Lys Ser Arg Pro  
100 105  
20 GAAAGCTCCG CAGAAATGAT TCCAAAACCC AGGGAGCAAC CACTGGCCCT ACCGTGGGAC 460  
TTACTCCCTC CTCTCCTTTG AGAGGCCCAT GTGTCGCTGG GGAGGAAGTG ACCCTTTGTG 520  
TAACTGTAAC CGAAAGTTTT TTCAAAAATC CTAGATGCTG TTGTTTGAAT GTTACATACT 580  
TCTATTTGTG CCACATCTCC CCTCCACTCC CCTGCTTAAT AAACCTCTAAA AATCCACTTG 640  
TATTTAATTC AGT 653

## (2) INFORMATION FOR SEQ ID NO: 44:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 439

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10408

135

## (ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 75.. 311  
 (C) CHARACTERIZATION METHOD: E

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GTAGAAACAG GCCTGTTAAG GAGAGGCCAC CGGGACTTCA GTGTCTCCTC CATCCCAGGA 60  
 GCGCAGTGGC CACT ATG GGG TCT GGG CTG CCC CTT GTC CTC CTC TTG ACC 110  
 10 Met Gly Ser Gly Leu Pro Leu Val Leu Leu Leu Thr  
                   1                  5                  10  
 CTC CTT GGC AGC TCA CAT GGA ACA GGG CCG GGT ATG ACT TTG CAA CTG 158  
 Leu Leu Gly Ser Ser His Gly Thr Gly Pro Gly Met Thr Leu Gln Leu  
                   15                  20                  25  
 15 AAG CTG AAG GAG TCT TTT CTG ACA AAT TCC TCC TAT GAG TCC AGC TTC 206  
 Lys Leu Lys Glu Ser Phe Leu Thr Asn Ser Ser Tyr Glu Ser Ser Phe  
                   30                  35                  40  
 CTG GAA TTG CTT GAA AAG CTC TGC CTC CTC CTC CAT CTC CCT TCA GGG 254  
 Leu Glu Leu Leu Glu Lys Leu Cys Leu Leu Leu His Leu Pro Ser Gly  
 20 45                  50                  55                  60  
 ACC AGC GTC ACC CTC CAC CAT GCA AGA TCT CAA CAC CAT GTT GTC TGC 302  
 Thr Ser Val Thr Leu His His Ala Arg Ser Gln His His Val Val Cys  
                   65                  70                  75  
 AAC ACA TGACAGCCAT TGAAGCCTGT GTCCTTCTTG GCCCGGGCTT TTGGGCCGGG GA 360  
 25 Asn Thr  
  
 TGCAGGAGGC AGGCCCCGAC CCTGTCTTTC AGCAGGCCCC CACCCTCCTG AGTGGCAATA 420  
 AATAAAATTC GGTATGCTG 439

30

## (2) INFORMATION FOR SEQ ID NO: 45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1131  
 (B) TYPE: Nucleic acid  
 35 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA



## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Stomach cancer  
 (D) CLONE NAME: HP10412

5

## (ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 56.. 1000  
 (C) CHARACTERIZATION METHOD: E

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

	CTATGAGATC CCGGCCTCAG GGTGGACGCA GTGGTTCTGC ACTGAGGCC TCGTC ATG	58
		Met
15		1
	GTG GCG CCT GTG TGG TAC TTG GTA GCG GCG GCT CTG CTA GTC GGC TTT	106
	Val Ala Pro Val Trp Tyr Leu Val Ala Ala Ala Leu Leu Val Gly Phe	
	5 10 15	
	ATC CTC TTC CTG ACT CGC AGC CGG GGC CGG GCG GCA TCA GCC GGC CAA	154
20	Ile Leu Phe Leu Thr Arg Ser Arg Gly Arg Ala Ala Ser Ala Gly Gln	
	20 25 30	
	GAG CCA CTG CAC AAT GAG GAG CTG GCA GGA GCA GGC CGG GTG GCC CAG	202
	Glu Pro Leu His Asn Glu Glu Leu Ala Gly Ala Gly Arg Val Ala Gln	
	35 40 45	
25	CCT GGG CCC CTG GAG CCT GAG GAG CCG AGA GCT GGA GGC AGG CCT CGG	250
	Pro Gly Pro Leu Glu Pro Glu Glu Pro Arg Ala Gly Gly Arg Pro Arg	
	50 55 60 65	
	CGC CGG AGG GAC CTG GGC AGC CGC CTA CAG GCC CAG CGT CGA GCC CAG	298
	Arg Arg Arg Asp Leu Gly Ser Arg Leu Gln Ala Gln Arg Arg Ala Gln	
30	70 75 80	
	CGG GTG GCC TGG GCA GAA GCA GAT GAG AAC GAG GAG GAA GCT GTC ATC	346
	Arg Val Ala Trp Ala Glu Ala Asp Glu Asn Glu Glu Glu Ala Val Ile	
	85 90 95	
	CTA GCC CAG GAG GAG GAA GGT GTC GAG AAG CCA GCG GAA ACT CAC CTG	394
35	Leu Ala Gln Glu Glu Glu Gly Val Glu Lys Pro Ala Glu Thr His Leu	
	100 105 110	
	TCG GGG AAA ATT GGA GCT AAG AAA CTG CGG AAG CTG GAG GAG AAA CAA	442
	Ser Gly Lys Ile Gly Ala Lys Lys Leu Arg Lys Leu Glu Glu Lys Gln	

0944555-12049

137

	115	120	125	
	GCG CGA AAG GCC CAG CGT GAG GCA GAG GAG GCT GAA CGT GAG GAG CGG			490
	Ala Arg Lys Ala Gln Arg Glu Ala Glu Glu Ala Glu Arg Glu Glu Arg			
	130	135	140	145
5	AAA CGA CTC GAG TCC CAG CGC GAA GCT GAG TGG AAG AAG GAG GAG GAG			538
	Lys Arg Leu Glu Ser Gln Arg Glu Ala Glu Trp Lys Lys Glu Glu Glu			
	150	155	160	
	CGG CTT CGC CTG GAG GAG GAG CAG AAG GAG GAG GAG GAG AGG AAG GCC			586
	Arg Leu Arg Leu Glu Glu Glu Gln Lys Glu Glu Glu Glu Arg Lys Ala			
10	165	170	175	
	CGC GAG GAG CAG GCC CAG CGG GAG CAT GAG GAG TAC CTG AAA CTG AAG			634
	Arg Glu Glu Gln Ala Gln Arg Glu His Glu Glu Tyr Leu Lys Leu Lys			
	180	185	190	
	GAG GCC TTT GTG GTG GAG GAG GAA GGC GTA GGA GAG ACC ATG ACT GAG			682
15	Glu Ala Phe Val Val Glu Glu Glu Gly Val Gly Glu Thr Met Thr Glu			
	195	200	205	
	GAA CAG TCC CAG AGC TTC CTG ACA GAG TTC ATC AAC TAC ATC AAG CAG			730
	Glu Gln Ser Gln Ser Phe Leu Thr Glu Phe Ile Asn Tyr Ile Lys Gln			
	210	215	220	225
20	TCC AAG GTT GTG CTC TTG GAA GAC CTG GCT TCC CAG GTG GGC CTA CGC			778
	Ser Lys Val Val Leu Leu Glu Asp Leu Ala Ser Gln Val Gly Leu Arg			
	230	235	240	
	ACT CAG GAC ACC ATA AAT CGC ATC CAG GAC CTG CTG GCT GAG GGG ACT			826
	Thr Gln Asp Thr Ile Asn Arg Ile Gln Asp Leu Leu Ala Glu Gly Thr			
25	245	250	255	
	ATA ACA GGT GTG ATT GAC GAC CGG GGC AAG TTC ATC TAC ATA ACC CCA			874
	Ile Thr Gly Val Ile Asp Asp Arg Gly Lys Phe Ile Tyr Ile Thr Pro			
	260	265	270	
	GAG GAA CTG GCC GCC GTG GCC AAC TTC ATC CGA CAG CGG GGC CGG GTG			922
30	Glu Glu Leu Ala Ala Val Ala Asn Phe Ile Arg Gln Arg Gly Arg Val			
	275	280	285	
	TCC ATC GCC GAG CTT GCC CAA GCC AGC AAC TCC CTC ATC GCC TGG GGC			970
	Ser Ile Ala Glu Leu Ala Gln Ala Ser Asn Ser Leu Ile Ala Trp Gly			
	290	295	300	305
35	CGG GAG TCC CCT GCC CAA GCC CCA GCC TGACCCAGT CCTTCCCTCT TGG			1020
	Arg Glu Ser Pro Ala Gln Ala Pro Ala			
	310			
	ACTCAGAGTT GGTGTGGCCT ACCTGGCTAT ACATCTTCAT CCCTCCCCAC CATCCTGGGG			1080

0044558-12019

AAGTGATGGT GTGGCCAGGC AGTTATAGAT TAAAGGCCTG TGAGTACTGC T

1131

## (2) INFORMATION FOR SEQ ID NO: 46:

## 5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1875  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 10 (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Stomach cancer  
 (D) CLONE NAME: HP10413

## (ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 79.. 666  
 (C) CHARACTERIZATION METHOD: E

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

CTCGCTCGCT CAGAGGGAGG AGAAAGTGGC GAGTTCCGGA TCCCTGCCTA GCGCGGCCCA 60  
 25 ACCTTTACTC CAGAGATC ATG GCT GCC GAG GAT GTG GTG GCG ACT GGC GCC 111  
                   Met Ala Ala Glu Asp Val Val Ala Thr Gly Ala  
                           1                          5                          10  
 GAC CCA AGC GAT CTG GAG AGC GGC GGG CTG CTG CAT GAG ATT TTC ACG 159  
 Asp Pro Ser Asp Leu Glu Ser Gly Gly Leu Leu His Glu Ile Phe Thr  
 30                  15                          20                          25  
 TCG CCG CTC AAC CTG CTG CTG CTT GGC CTC TGC ATC TTC CTG CTC TAC 207  
 Ser Pro Leu Asn Leu Leu Leu Leu Gly Leu Cys Ile Phe Leu Leu Tyr  
                   30                          35                          40  
 AAG ATC GTG CGC GGC GAC CAG CCG GCG GCC AGC GGC GAC AGC GAC GAC 255  
 35 Lys Ile Val Arg Gly Asp Gln Pro Ala Ala Ser Gly Asp Ser Asp Asp  
                   45                          50                          55  
 GAC GAG CCG CCC CCT CTG CCC CGC CTC AAG CGG CGC GAC TTC ACC CCC 303  
 Asp Glu Pro Pro Pro Leu Pro Arg Leu Lys Arg Arg Asp Phe Thr Pro

0944558.120199

139

	60	65	70	75	
	GCC GAG CTG CGG CGC TTC GAC GGC GTC CAG GAC CCG CGC ATA CTC ATG				351
	Ala Glu Leu Arg Arg Phe Asp Gly Val Gln Asp Pro Arg Ile Leu Met				
		80	85	90	
5	GCC ATC AAC GGC AAG GTG TTC GAT GTG ACC AAA GGC CGC AAA TTC TAC				399
	Ala Ile Asn Gly Lys Val Phe Asp Val Thr Lys Gly Arg Lys Phe Tyr				
		95	100	105	
	GGG CCC GAG GGC CCG TAT GGC GTC TTT GCT GGA AGA GAT GCA TCC AGG				447
	Gly Pro Glu Gly Pro Tyr Gly Val Phe Ala Gly Arg Asp Ala Ser Arg				
10		110	115	120	
	GGC CTT GCC ACA TTT TGC CTG GAT AAG GAA GCA CTG AAG GAT GAG TAC				495
	Gly Leu Ala Thr Phe Cys Leu Asp Lys Glu Ala Leu Lys Asp Glu Tyr				
		125	130	135	
	GAT GAC CTT TCT GAC CTC ACT GCT GCC CAG CAG GAG ACT CTG AGT GAC				543
15	Asp Asp Leu Ser Asp Leu Thr Ala Ala Gln Gln Glu Thr Leu Ser Asp				
		140	145	150	155
	TGG GAG TCT CAG TTC ACT TTC AAG TAT CAT CAC GTG GGC AAA CTG CTG				591
	Trp Glu Ser Gln Phe Thr Phe Lys Tyr His His Val Gly Lys Leu Leu				
		160	165	170	
20	AAG GAG GGC GAG GAG CCC ACT GTG TAC TCA GAT GAG GAA GAA CCA AAA				639
	Lys Glu Gly Glu Glu Pro Thr Val Tyr Ser Asp Glu Glu Glu Pro Lys				
		175	180	185	
	GAT GAG AGT GCC CGG AAA AAT GAT TAAAGCATTC AGTGGAAGTA TATCTAT				690
	Asp Glu Ser Ala Arg Lys Asn Asp				
25		190	195		
	TTTTGTATTT TGCAAAATCA TTTGTAACAG TCCACTCTGT CTTTAAAACA TAGTGATTAC				750
	AATATTTAGA AAGTTTTGAG CACTTGCTAT AAGTTTTTTA TAACATCACT AGTGACACTA				810
	ATAAAATTAA CTTCTTAGAA TGCATGATGT GTTTGTGTGT CACAAATCCA GAAAGTGAAC				870
	TGCAGTGCTG TAATACACAT GTTAATACTG TTTTCTTCT ATCTGTAGTT AGTACAGGAT				930
30	GAATTTAAAT GTGTTTTTCC TGAGAGACAA GGAAGACTTG GGTATTTCCC AAAACAGGTA				990
	AAAATCTTAA ATGTGCACCA AGAGCAAAGG ATCAACTTTT AGTCATGATG TTCTGTAAAG				1050
	ACAACAAATC CCTTTTTTTT TCTCAATTGA CTTAACTGCA TGATTTCTGT TTTATCTACC				1110
	TCTAAAGCAA ATCTGCAGTG TTCCAAAGAC TTTGGTATGG ATTAAGCGCT GTCCAGTAAC				1170
	AAAATGAAAT CTCAAAACAG AGCTCAGCTG CAAAAAAGCA TATTTTCTGT GTTTCTGGAC				1230
35	TGCACTGTTG TCCTTGCCCT CACATAGACA CTCAGACACC CTCACAAACA CAGTAGTCTA				1290
	TAGTTAGGAT TAAAATAGGA TCTGAACATT CAAAAGAAAG CTTTGAAAAA AAAGAGCTGG				1350
	CTGGCCTAAA AACCTAAATA TATGATGAAG ATTGTAGGAC TGTCTTCCCA AGCCCCATGT				1410
	TCATGGTGGG GCAATGGTTA TTTGGTTATT TTA CTCAATT GGTACTCTC ATTTGAAATG				1470

66T02"852460

140

AGGGAGGGAC ATACAGAATA GGAACAGGTG TTTGCTCTCC TAAGAGCCTT CATGCACACC 1530  
 CCTGAACCAC GAGGAAACAG TACAGTCGCT AGTCAAGTGG TTTTAAAGT AAAGTATATT 1590  
 CATAAGGTAA CAGTTATTCT GTTGTTATAA AACTATACCC ACTGCAAAAG TAGTAGTCAA 1650  
 GTGTCTAGGT CTTTGATATT GCTCTTTTGG TTAACACTAA GCTTAAGTAG ACTATACAGT 1710  
 5 TGTATGAATT TGTAAGAGTA TATGAACACC TAGTGAGATT TCAAACTTGT AATTGTGGTT 1770  
 AAATAGTCAT TGTATTTTCT TGTGAACTGT GTTTTATGAT TTTACCTCAA ATCAGAAAAC 1830  
 AAAATGATGT GCTTTGGTCA GTTAATAAAA ATGGTTTTAC CCACT 1875

10 (2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1563

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

15 (D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

20 (B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10415

(ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

25 (B) EXISTENCE POSITION: 72.. 1460

(C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

30 AAATTGGGCC AGGCTGAGGC GCTGCTGCTG GAGCGGCCGA TCCGAGACGT GGCTCCCTGG 60  
 GCGGCAGAAC C ATG TTG GAC TTC GCG ATC TTC GCC GTT ACC TTC TTG CTG 110  
 Met Leu Asp Phe Ala Ile Phe Ala Val Thr Phe Leu Leu  
 1 5 10  
 GCG TTG GTG GGA GCC GTG CTC TAC CTC TAT CCG GCT TCC AGA CAA GCT 158  
 35 Ala Leu Val Gly Ala Val Leu Tyr Leu Tyr Pro Ala Ser Arg Gln Ala  
 15 20 25  
 GCA GGA ATT CCA GGG ATT ACT CCA ACT GAA GAA AAA GAT GGT AAT CTT 206  
 Ala Gly Ile Pro Gly Ile Thr Pro Thr Glu Glu Lys Asp Gly Asn Leu

09445558 120199

141

	30		35		40		45										
	CCA	GAT	ATT	GTG	AAT	AGT	GGA	AGT	TTG	CAT	GAG	TTC	CTG	GTT	AAT	TTG	254
	Pro	Asp	Ile	Val	Asn	Ser	Gly	Ser	Leu	His	Glu	Phe	Leu	Val	Asn	Leu	
				50					55						60		
5	CAT	GAG	AGA	TAT	GGG	CCT	GTG	GTC	TCC	TTC	TGG	TTT	GGC	AGG	CGC	CTC	302
	His	Glu	Arg	Tyr	Gly	Pro	Val	Val	Ser	Phe	Trp	Phe	Gly	Arg	Arg	Leu	
				65					70						75		
	GTG	GTT	AGT	TTG	GGC	ACT	GTT	GAT	GTA	CTG	AAG	CAG	CAT	ATC	AAT	CCC	350
	Val	Val	Ser	Leu	Gly	Thr	Val	Asp	Val	Leu	Lys	Gln	His	Ile	Asn	Pro	
10			80						85						90		
	AAT	AAG	ACA	TTG	GAC	CCT	TTT	GAA	ACC	ATG	CTG	AAG	TCA	TTA	TTA	AGG	398
	Asn	Lys	Thr	Leu	Asp	Pro	Phe	Glu	Thr	Met	Leu	Lys	Ser	Leu	Leu	Arg	
			95						100						105		
	TAT	CAA	TCT	GGT	GGT	GGC	AGT	GTG	AGT	GAA	AAC	CAC	ATG	AGG	AAA	AAA	446
15	Tyr	Gln	Ser	Gly	Gly	Gly	Ser	Val	Ser	Glu	Asn	His	Met	Arg	Lys	Lys	
	110					115					120					125	
	TTG	TAT	GAA	AAT	GGT	GTG	ACT	GAT	TCT	CTG	AAG	AGT	AAC	TTT	GCC	CTC	494
	Leu	Tyr	Glu	Asn	Gly	Val	Thr	Asp	Ser	Leu	Lys	Ser	Asn	Phe	Ala	Leu	
				130						135					140		
20	CTC	CTA	AAG	CTT	TCA	GAA	GAA	TTA	TTA	GAT	AAA	TGG	CTC	TCC	TAC	CCA	542
	Leu	Leu	Lys	Leu	Ser	Glu	Glu	Leu	Leu	Asp	Lys	Trp	Leu	Ser	Tyr	Pro	
				145					150						155		
	GAG	ACC	CAG	CAC	GTG	CCC	CTC	AGC	CAG	CAT	ATG	CTT	GGT	TTT	GCT	ATG	590
	Glu	Thr	Gln	His	Val	Pro	Leu	Ser	Gln	His	Met	Leu	Gly	Phe	Ala	Met	
25			160						165						170		
	AAG	TCT	GTT	ACA	CAG	ATG	GTA	ATG	GGT	AGT	ACA	TTT	GAA	GAT	GAT	CAG	638
	Lys	Ser	Val	Thr	Gln	Met	Val	Met	Gly	Ser	Thr	Phe	Glu	Asp	Asp	Gln	
			175						180						185		
	GAA	GTC	ATT	CGC	TTC	CAG	AAG	AAT	CAT	GGC	ACA	GTT	TGG	TCT	GAG	ATT	686
30	Glu	Val	Ile	Arg	Phe	Gln	Lys	Asn	His	Gly	Thr	Val	Trp	Ser	Glu	Ile	
	190					195					200				205		
	GGA	AAA	GGC	TTT	CTA	GAT	GGG	TCA	CTT	GAT	AAA	AAC	ATG	ACT	CGG	AAA	734
	Gly	Lys	Gly	Phe	Leu	Asp	Gly	Ser	Leu	Asp	Lys	Asn	Met	Thr	Arg	Lys	
				210						215					220		
35	AAA	CAA	TAT	GAA	GAT	GCC	CTC	ATG	CAA	CTG	GAG	TCT	GTT	TTA	AGG	AAC	782
	Lys	Gln	Tyr	Glu	Asp	Ala	Leu	Met	Gln	Leu	Glu	Ser	Val	Leu	Arg	Asn	
				225						230					235		
	ATC	ATA	AAA	GAA	CGA	AAA	GGA	AGG	AAC	TTC	AGT	CAA	CAT	ATT	TTC	ATT	830

0944353 120199

142

	Ile	Ile	Lys	Glu	Arg	Lys	Gly	Arg	Asn	Phe	Ser	Gln	His	Ile	Phe	Ile	
	240				245				250								
	GAC	TCC	TTA	GTA	CAA	GGG	AAC	CTT	AAT	GAC	CAA	CAG	ATC	CTA	GAA	GAC	878
	Asp	Ser	Leu	Val	Gln	Gly	Asn	Leu	Asn	Asp	Gln	Gln	Ile	Leu	Glu	Asp	
5	255				260				265								
	AGT	ATG	ATA	TTT	TCT	CTG	GCC	AGT	TGC	ATA	ATA	ACT	GCA	AAA	TTG	TGT	926
	Ser	Met	Ile	Phe	Ser	Leu	Ala	Ser	Cys	Ile	Ile	Thr	Ala	Lys	Leu	Cys	
	270				275				280				285				
	ACC	TGG	GCA	ATC	TGT	TTT	TTA	ACC	ACC	TCT	GAA	GAA	GTT	CAA	AAA	AAA	974
10	Thr	Trp	Ala	Ile	Cys	Phe	Leu	Thr	Thr	Ser	Glu	Glu	Val	Gln	Lys	Lys	
	290				295				300								
	TTA	TAT	GAA	GAG	ATA	AAC	CAA	GTT	TTT	GGA	AAT	GGT	CCT	GTT	ACT	CCA	1022
	Leu	Tyr	Glu	Glu	Ile	Asn	Gln	Val	Phe	Gly	Asn	Gly	Pro	Val	Thr	Pro	
	305				310				315								
15	GAG	AAA	ATT	GAG	CAG	CTC	AGA	TAT	TGT	CAG	CAT	GTG	CTT	TGT	GAA	ACT	1070
	Glu	Lys	Ile	Glu	Gln	Leu	Arg	Tyr	Cys	Gln	His	Val	Leu	Cys	Glu	Thr	
	320				325				330								
	GTT	CGA	ACT	GCC	AAA	CTG	ACT	CCA	GTT	TCT	GCC	CAG	CTT	CAA	GAT	ATT	1118
	Val	Arg	Thr	Ala	Lys	Leu	Thr	Pro	Val	Ser	Ala	Gln	Leu	Gln	Asp	Ile	
20	335				340				345								
	GAA	GGA	AAA	ATT	GAC	CGA	TTT	ATT	ATT	CCT	AGA	GAG	ACC	CTC	GTC	CTT	1166
	Glu	Gly	Lys	Ile	Asp	Arg	Phe	Ile	Ile	Pro	Arg	Glu	Thr	Leu	Val	Leu	
	350				355				360				365				
	TAT	GCC	CTT	GGT	GTG	GTA	CTT	CAG	GAT	CCT	AAT	ACT	TGG	CCA	TCT	CCA	1214
25	Tyr	Ala	Leu	Gly	Val	Val	Leu	Gln	Asp	Pro	Asn	Thr	Trp	Pro	Ser	Pro	
	370				375				380								
	CAC	AAG	TTT	GAT	CCA	GAT	CGG	TTT	GAT	GAT	GAA	TTA	GTA	ATG	AAA	ACT	1262
	His	Lys	Phe	Asp	Pro	Asp	Arg	Phe	Asp	Asp	Glu	Leu	Val	Met	Lys	Thr	
	385				390				395								
30	TTT	TCC	TCA	CTT	GGA	TTC	TCA	GGC	ACA	CAG	GAG	TGT	CCA	GAG	TTG	AGG	1310
	Phe	Ser	Ser	Leu	Gly	Phe	Ser	Gly	Thr	Gln	Glu	Cys	Pro	Glu	Leu	Arg	
	400				405				410								
	TTT	GCA	TAT	ATG	GTG	ACC	ACA	GTA	CTT	CTT	AGT	GTA	TTG	GTG	AAG	AGA	1358
	Phe	Ala	Tyr	Met	Val	Thr	Thr	Val	Leu	Leu	Ser	Val	Leu	Val	Lys	Arg	
35	415				420				425								
	CTG	CAC	CTA	CTT	TCT	GTG	GAG	GGA	CAG	GTT	ATT	GAA	ACA	AAG	TAT	GAA	1406
	Leu	His	Leu	Leu	Ser	Val	Glu	Gly	Gln	Val	Ile	Glu	Thr	Lys	Tyr	Glu	
	430				435				440				445				

66T02T 352460

143

CTG GTA ACA TCA TCA AGG GAA GAA GCT TGG ATC ACT GTC TCA AAG AGA 1454  
 Leu Val Thr Ser Ser Arg Glu Glu Ala Trp Ile Thr Val Ser Lys Arg  
 450 455 460

TAT TAAAATTTTA TACATTTAAA ATCATTGTGA AATTGATTGA GGAAAACAAC CAT 1510

5 Tyr

TTAAAAA TCTATGTTGA ATCCTTTTAT AAACCAGTAT CACTTTGTAA TAT 1563

10 (2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2030  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Stomach cancer  
 (D) CLONE NAME: HP10419

(ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 171.. 914  
 (C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

30 CATTTGGGGT TTCGGTTCCT CCCCTTCCCC TTCCCCGGGG TCTGGGGGTG ACATTGCACC 60  
 GCGCCCCTCG TGGGGTCGCG TTGCCACCCC ACGCGGACTC CCCAGCTGGC GCGCCCCTCC 120  
 CATTTGCCTG TCCTGGTCAG GCCCCACCC CCCTTCCCAC CTGACCAGCC ATG GGG 176  
 Met Gly  
 1  
 35 GCT GCG GTG TTT TTC GGC TGC ACT TTC GTC GCG TTC GGC CCG GCC TTC 224  
 Ala Ala Val Phe Phe Gly Cys Thr Phe Val Ala Phe Gly Pro Ala Phe  
 5 10 15  
 GCG CTT TTC TTG ATC ACT GTG GCT GGG GAC CCG CTT CGC GTT ATC ATC 272

044558 12019  
 66T02 8525460



144

	Ala	Leu	Phe	Leu	Ile	Thr	Val	Ala	Gly	Asp	Pro	Leu	Arg	Val	Ile	Ile	
	20						25					30					
	CTG	GTC	GCA	GGG	GCA	TTT	TTC	TGG	CTG	GTC	TCC	CTG	CTC	CTG	GCC	TCT	320
	Leu	Val	Ala	Gly	Ala	Phe	Phe	Trp	Leu	Val	Ser	Leu	Leu	Leu	Ala	Ser	
5	35					40					45					50	
	GTG	GTC	TGG	TTC	ATC	TTG	GTC	CAT	GTG	ACC	GAC	CGG	TCA	GAT	GCC	CGG	368
	Val	Val	Trp	Phe	Ile	Leu	Val	His	Val	Thr	Asp	Arg	Ser	Asp	Ala	Arg	
					55					60					65		
	CTC	CAG	TAC	GGC	CTC	CTG	ATT	TTT	GGT	GCT	GCT	GTC	TCT	GTC	CTT	CTA	416
10	Leu	Gln	Tyr	Gly	Leu	Leu	Ile	Phe	Gly	Ala	Ala	Val	Ser	Val	Leu	Leu	
				70					75						80		
	CAG	GAG	GTG	TTC	CGC	TTT	GCC	TAC	TAC	AAG	CTG	CTT	AAG	AAG	GCA	GAT	464
	Gln	Glu	Val	Phe	Arg	Phe	Ala	Tyr	Tyr	Lys	Leu	Leu	Lys	Lys	Ala	Asp	
				85			90						95				
15	GAG	GGG	TTA	GCA	TCG	CTG	AGT	GAG	GAC	GGA	AGA	TCA	CCC	ATC	TCC	ATC	512
	Glu	Gly	Leu	Ala	Ser	Leu	Ser	Glu	Asp	Gly	Arg	Ser	Pro	Ile	Ser	Ile	
				100			105						110				
	CGC	CAG	ATG	GCC	TAT	GTT	TCT	GGT	CTC	TCC	TTC	GGT	ATC	ATC	AGT	GGT	560
	Arg	Gln	Met	Ala	Tyr	Val	Ser	Gly	Leu	Ser	Phe	Gly	Ile	Ile	Ser	Gly	
20	115					120					125				130		
	GTC	TTC	TCT	GTT	ATC	AAT	ATT	TTG	GCT	GAT	GCA	CTT	GGG	CCA	GGT	GTG	608
	Val	Phe	Ser	Val	Ile	Asn	Ile	Leu	Ala	Asp	Ala	Leu	Gly	Pro	Gly	Val	
					135					140					145		
	GTT	GGG	ATC	CAT	GGA	GAC	TCA	CCC	TAT	TAC	TTC	CTG	ACT	TCA	GCC	TTT	656
25	Val	Gly	Ile	His	Gly	Asp	Ser	Pro	Tyr	Tyr	Phe	Leu	Thr	Ser	Ala	Phe	
				150						155					160		
	CTG	ACA	GCA	GCC	ATT	ATC	CTG	CTC	CAT	ACC	TTT	TGG	GGA	GTT	GTG	TTC	704
	Leu	Thr	Ala	Ala	Ile	Ile	Leu	Leu	His	Thr	Phe	Trp	Gly	Val	Val	Phe	
				165					170						175		
30	TTT	GAT	GCC	TGT	GAG	AGG	AGA	CGG	TAC	TGG	GCT	TTG	GGC	CTG	GTG	GTT	752
	Phe	Asp	Ala	Cys	Glu	Arg	Arg	Arg	Tyr	Trp	Ala	Leu	Gly	Leu	Val	Val	
				180					185						190		
	GGG	AGT	CAC	CTA	CTG	ACA	TCG	GGA	CTG	ACA	TTC	CTG	AAC	CCC	TGG	TAT	800
	Gly	Ser	His	Leu	Leu	Thr	Ser	Gly	Leu	Thr	Phe	Leu	Asn	Pro	Trp	Tyr	
35	195					200					205				210		
	GAG	GCC	AGC	CTG	CTG	CCC	ATC	TAT	GCA	GTC	ACT	GTT	TCC	ATG	GGG	CTC	848
	Glu	Ala	Ser	Leu	Leu	Pro	Ile	Tyr	Ala	Val	Thr	Val	Ser	Met	Gly	Leu	
					215						220				225		

0944558 120159

145

TGG GCC TTC ATC ACA GCT GGA GGG TCC CTC CGA AGT ATT CAG CGC AGC 896  
Trp Ala Phe Ile Thr Ala Gly Gly Ser Leu Arg Ser Ile Gln Arg Ser

230

235

240

CTC TTG TGT AAG GAC TGACTACCTG GACTGATCGC CTGACAGATC CCACCTGCC 950

5 Leu Leu Cys Lys Asp

245

TGTCCACTGC CCATGACTGA GCCCAGCCCC AGCCCGGGTC CATTGCCCAC ATTCTCTGTC 1010

TCCTTCTCGT CGGTCTACCC CACTACCTCC AGGGTTTTGC TTTGTCCTTT TGTGACCGTT 1070

AGTCTCTAAG CTTTACCAGG AGCAGCCTGG GTTCAGCCAG TCAGTGA CTG GTGGGTTTGA 1130

10 ATCTGCACTT ATCCCCACCA CCTGGGGACC CCCTTGTTGT GTCCAGGACT CCCCCTGTGT 1190

CAGTGCTCTG CTCTCACCCT GCCCAAGACT CACCTCCCTT CCCCTCTGCA GGCCGACGGC 1250

AGGAGGACAG TCGGGTGATG GTGTATTCTG CCCTGCGCAT CCCACCCGAG GACTGAGGGA 1310

ACCTAGGGGG GACCCCTGGG CCTGGGGTGC CCTCCTGATG TCCTCGCCCT GTATTTCTCC 1370

ATCTCCAGTT CTGGACAGTG CAGGTTGCCA AGAAAAGGGA CCTAGTTTAG CCATTGCCCT 1430

15 GGAGATGAAA TTAATGGAGG CTCAAGGATA GATGAGCTCT GAGTTTCTCA GTACTCCCTC 1490

AAGACTGGAC ATCTTGGTCT TTTTCTCAGG CCTGAGGGGG AACCATTTTT GGTGTGATAA 1550

ATACCCTAAA CTGCCTTTTT TTCTTTTTTG AGGTGGGGGG AGGGAGGAGG TATATTGGAA 1610

CTCTTCTAAC CTCCTTGGGC TATATTTTCT CTCCTCGAGT TGCTCCTCAT GGCTGGGCTC 1670

ATTTTCGGTCC CTTTCTCCTT GGTCCCAGAC CTTGGGGGAA AGGAAGGAAG TGCATGTTTG 1730

20 GGAAGTGGCA TTAAGTGAAC TAATGGTTTT AACCTCCTTA ACCACCAGCA TCCCTCCTCT 1790

CCCCAAGGTG AAGTGGAGGG TGCTGTGGTG AGCTGGCCAC TCCAGAGCTG CAGTGCCACT 1850

GGAGGAGTCA GACTACCATG ACATCGTAGG GAAGGAGGGG AGATTTTTTT GTAGTTTTTA 1910

ATTGGGGTGT GGGAGGGGCG GGGAGTTTTT CTATAAACTG TATCATTTTC TGCTGAGGGT 1970

GGAGTGTCCC ATCCTTTTAA TCAAGGTGAT TGTGATTTTG ACTAATAAAA AAGAATTTGT 2030

25

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 493

30 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

35 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Stomach cancer

(D) CLONE NAME: HP10424

146

## (ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 98.. 439

(C) CHARACTERIZATION METHOD: E

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

AAAGTTTCCC AAATCCAGGC GGCTAGAGGC CCACTGCTTC CCAACTACCA GCTGAGGGGG 60  
 TCCGTCCCGA GAAGGGAGAA GAGGCCGAAG AGGAAAC ATG AAC TTC TAT TTA CTC 115  
 10 Met Asn Phe Tyr Leu Leu  
 1 5  
 CTA GCG AGC AGC ATT CTG TGT GCC TTG ATT GTC TTC TGG AAA TAT CGC 163  
 Leu Ala Ser Ser Ile Leu Cys Ala Leu Ile Val Phe Trp Lys Tyr Arg  
 10 15 20  
 15 CGC TTT CAG AGA AAC ACT GGC GAA ATG TCA TCA AAT TCA ACT GCT CTT 211  
 Arg Phe Gln Arg Asn Thr Gly Glu Met Ser Ser Asn Ser Thr Ala Leu  
 25 30 35  
 GCA CTA GTG AGA CCC TCT TCT TCT GGG TTA ATT AAC AGC AAT ACA GAC 259  
 Ala Leu Val Arg Pro Ser Ser Ser Gly Leu Ile Asn Ser Asn Thr Asp  
 20 40 45 50  
 AAC AAT CTT GCA GTC TAC GAC CTC TCT CGG GAT ATT TTA AAT AAT TTC 307  
 Asn Asn Leu Ala Val Tyr Asp Leu Ser Arg Asp Ile Leu Asn Asn Phe  
 55 60 65 70  
 CCA CAC TCA ATA GCC AGG CAG AAG CGA ATA TTG GTA AAC CTC AGT ATG 355  
 25 Pro His Ser Ile Ala Arg Gln Lys Arg Ile Leu Val Asn Leu Ser Met  
 75 80 85  
 GTG GAA AAC AAG CTG GTT GAA CTG GAA CAT ACT CTA CTT AGC AAG GGT 403  
 Val Glu Asn Lys Leu Val Glu Leu Glu His Thr Leu Leu Ser Lys Gly  
 90 95 100  
 30 TTC AGA GGT GCA TCA CCT CAC CGG AAA TCC ACC TAAAAGCGTA CAGG 450  
 Phe Arg Gly Ala Ser Pro His Arg Lys Ser Thr  
 105 110  
 ATGTAATGCC AGTGGTGGAA ATCATTAAAG ACACTTTGA GTAG 493

35

## (2) INFORMATION FOR SEQ ID NO: 50:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2044

09445258-120199

147

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Homo sapiens*

(B) CELL KIND: Epidermoid carcinoma

(C) CELL LINE: KB

10

(D) CLONE NAME: HP10428

(ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS

(B) EXISTENCE POSITION: 288.. 1385

15

(C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

AGATTCCGGC CTGGAGCTCC CAGGGCCGAG CAGACCTTGG GACCTGTGAG CGCTGCATCC 60  
 20 AATTAACCAT GGGAAGGGTC AGCACCAGCC ACCAGCCCCT TAGGTGAGGA CTCTGCCTGG 120  
 GGCTCTGCTG ATGGTTCCGA ATCATGGAGC TGCAGAGAGC TCCTCCAGCC TGGAGACGTT 180  
 CTTGGTGAAA GCTGTGGTCT AACTCCACCG GCTCTTCCTG CACATTGTAT TCAAGAGGGG 240  
 TGCCTGCCCC CGCTGACTCA GGAGCTCCGG TGCTGCAGCC GCCACGA ATG GGG AGG 296  
 Met Gly Arg  
 25 1  
 TGG GCC CTC GAT GTG GCC TTT TTG TGG AAG GCG GTG TTG ACC CTG GGG 344  
 Trp Ala Leu Asp Val Ala Phe Leu Trp Lys Ala Val Leu Thr Leu Gly  
 5 10 15  
 CTG GTG CTT CTC TAC TAC TGC TTC TCC ATC GGC ATC ACC TTC TAC AAC 392  
 30 Leu Val Leu Leu Tyr Tyr Cys Phe Ser Ile Gly Ile Thr Phe Tyr Asn  
 20 25 30 35  
 AAG TGG CTG ACA AAG AGC TTC CAT TTC CCC CTC TTC ATG ACG ATG CTG 440  
 Lys Trp Leu Thr Lys Ser Phe His Phe Pro Leu Phe Met Thr Met Leu  
 40 45 50  
 35 CAC CTG GCC GTG ATC TTC CTC TTC TCC GCC CTG TCC AGG GCG CTG GTT 488  
 His Leu Ala Val Ile Phe Leu Phe Ser Ala Leu Ser Arg Ala Leu Val  
 55 60 65  
 CAG TGC TCC AGC CAC AGG GCC CGT GTG GTG CTG AGC TGG GCC GAC TAC 536

004455508 10160

148

	Gln	Cys	Ser	Ser	His	Arg	Ala	Arg	Val	Val	Leu	Ser	Trp	Ala	Asp	Tyr	
	70					75					80						
	CTC	AGA	AGA	GTG	GCT	CCC	ACA	GCT	CTG	GCG	ACG	GCG	CTT	GAC	GTG	GGC	584
	Leu	Arg	Arg	Val	Ala	Pro	Thr	Ala	Leu	Ala	Thr	Ala	Leu	Asp	Val	Gly	
5	85					90					95						
	TTG	TCC	AAC	TGG	AGC	TTC	CTG	TAT	GTC	ACC	GTC	TCG	CTG	TAC	ACA	ATG	632
	Leu	Ser	Asn	Trp	Ser	Phe	Leu	Tyr	Val	Thr	Val	Ser	Leu	Tyr	Thr	Met	
	100					105					110					115	
	ACC	AAA	TCC	TCA	GCT	GTC	CTC	TTC	ATC	TTG	ATC	TTC	TCT	CTG	ATC	TTC	680
10	Thr	Lys	Ser	Ser	Ala	Val	Leu	Phe	Ile	Leu	Ile	Phe	Ser	Leu	Ile	Phe	
	120					125					130						
	AAG	CTG	GAG	GAG	CTG	CGC	GCG	GCA	CTG	GTC	CTG	GTG	GTC	CTC	CTC	ATC	728
	Lys	Leu	Glu	Glu	Leu	Arg	Ala	Ala	Leu	Val	Leu	Val	Val	Leu	Leu	Ile	
	135					140					145						
15	GCC	GGG	GGT	CTC	TTC	ATG	TTC	ACC	TAC	AAG	TCC	ACA	CAG	TTC	AAC	GTG	776
	Ala	Gly	Gly	Leu	Phe	Met	Phe	Thr	Tyr	Lys	Ser	Thr	Gln	Phe	Asn	Val	
	150					155					160						
	GAG	GGC	TTC	GCC	TTG	GTG	CTG	GGG	GCC	TCG	TTC	ATC	GGT	GGC	ATT	CGC	824
	Glu	Gly	Phe	Ala	Leu	Val	Leu	Gly	Ala	Ser	Phe	Ile	Gly	Gly	Ile	Arg	
20	165					170					175						
	TGG	ACC	CTC	ACC	CAG	ATG	CTC	CTG	CAG	AAG	GCT	GAA	CTC	GGC	CTC	CAG	872
	Trp	Thr	Leu	Thr	Gln	Met	Leu	Leu	Gln	Lys	Ala	Glu	Leu	Gly	Leu	Gln	
	180					185					190					195	
	AAT	CCC	ATC	GAC	ACC	ATG	TTC	CAC	CTG	CAG	CCA	CTC	ATG	TTC	CTG	GGG	920
25	Asn	Pro	Ile	Asp	Thr	Met	Phe	His	Leu	Gln	Pro	Leu	Met	Phe	Leu	Gly	
	200					205					210						
	CTC	TTC	CCT	CTC	TTT	GCT	GTA	TTT	GAA	GGT	CTC	CAT	TTG	TCC	ACA	TCT	968
	Leu	Phe	Pro	Leu	Phe	Ala	Val	Phe	Glu	Gly	Leu	His	Leu	Ser	Thr	Ser	
	215					220					225						
30	GAG	AAA	ATC	TTC	CGT	TTC	CAG	GAC	ACA	GGG	CTG	CTC	CTG	CGG	GTA	CTT	1016
	Glu	Lys	Ile	Phe	Arg	Phe	Gln	Asp	Thr	Gly	Leu	Leu	Leu	Arg	Val	Leu	
	230					235					240						
	GGG	AGC	CTC	TTC	CTT	GGC	GGG	ATT	CTC	GCC	TTT	GGT	TTG	GGC	TTC	TCT	1064
	Gly	Ser	Leu	Phe	Leu	Gly	Gly	Ile	Leu	Ala	Phe	Gly	Leu	Gly	Phe	Ser	
35	245					250					255						
	GAG	TTC	CTC	CTG	GTC	TCC	AGA	ACC	TCC	AGC	CTC	ACT	CTC	TCC	ATT	GCC	1112
	Glu	Phe	Leu	Leu	Val	Ser	Arg	Thr	Ser	Ser	Leu	Thr	Leu	Ser	Ile	Ala	
	260					265					270					275	

0944558 120199

149

GGC ATT TTT AAG GAA GTC TGC ACT TTG CTG TTG GCA GCT CAT CTG CTG 1160  
 Gly Ile Phe Lys Glu Val Cys Thr Leu Leu Leu Ala Ala His Leu Leu  
 280 285 290

GGC GAT CAG ATC AGC CTC CTG AAC TGG CTG GGC TTC GCC CTC TGC CTC 1208  
 5 Gly Asp Gln Ile Ser Leu Leu Asn Trp Leu Gly Phe Ala Leu Cys Leu  
 295 300 305

TCG GGA ATA TCC CTC CAC GTT GCC CTC AAA GCC CTG CAT TCC AGA GGT 1256  
 Ser Gly Ile Ser Leu His Val Ala Leu Lys Ala Leu His Ser Arg Gly  
 310 315 320

10 GAT GGT GGC CCC AAG GCC TTG AAG GGG CTG GGC TCC AGC CCC GAC CTG 1304  
 Asp Gly Gly Pro Lys Ala Leu Lys Gly Leu Gly Ser Ser Pro Asp Leu  
 325 330 335

GAG CTG CTG CTC CGG AGC AGC CAG CGG GAG GAA GGT GAC AAT GAG GAG 1352  
 Glu Leu Leu Leu Arg Ser Ser Gln Arg Glu Glu Gly Asp Asn Glu Glu  
 15 340 345 350 355

GAG GAG TAC TTT GTG GCC CAG GGG CAG CAG TGACCAGCCA GGGCAAAT 1400  
 Glu Glu Tyr Phe Val Ala Gln Gly Gln Gln  
 360 365

GGCTTAGAAG CAGGCCACTC CCCAGCCTGC TGCCAGCACT CACTGTGCTC AAGCCGCCAG 1460  
 20 GGCTCATCAT GGTAGCTGGG AGCTGTGGAC GGGAGTCACC AGGTGGTGGG GCCAAGCCAG 1520  
 GGA CTG CAT CTA TTTTGGCCCC TCCCTTCAGA GCCTGGTCAC ACAAGGGGCG AGCACCAGGC 1580  
 CAGCCTGGGA CTGGCCAGAG CTGGGCCCAA GCTGCGCTGG AATCGCAGCA GGAGAGGGGA 1640  
 GTGGGCTGGT TCTTCCCACC ACTTCCCAGG CTCTGACAGC CGAGACTCAT TTCCAAGGCA 1700  
 CAGCAGCTTT CTAAAGGGAC TGAGTTTGGA CTGGGTTTTG GACCTCCAGG GGCTGGAGCT 1760

25 TCATCACCTG GGCAGTGTCT TTTCTCAGAG AGCAGGTTTC TTTATAGTTT GAAATAAAT 1820  
 GGTTCACGGT CCACTGGCCG CTTGTGTGTTG CTGGAGACGT GGGGGCAGGG AGGGGACAGT 1880  
 GTGGGCCTGG CCTCTCCTTT CCTTCCCTG CCTGGAGCCT TCTTCAAATG TCTGGTCTTA 1940  
 AGCCAGGCCT CCTTCATTTT CTCGCTCCTG TTAGAACACC AGTCCCCTCC CCAGTGGGGC 2000  
 CCCACTGCAC CTGCTGGCAG GAAATAAATG AATGTTTACT GAGT 2044

30

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1043

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

(A) ORGANISM: *Homo sapiens*  
(B) CELL KIND: Stomach cancer  
(D) CLONE NAME: HP10429

(ix) SEQUENCE CHARACTERISTICS:

(A) CHARACTERIZATION CODE: CDS  
(B) EXISTENCE POSITION: 157.. 837  
(C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

ATTAGCATAA	CCCTTCCTCA	GGAAGAGTGA	GATTTTATAT	TTGACAATAA	AGTGTTAGAC	60
TCCATTTCTA	AATACCAGAC	TTCAAAAGAT	AAGGTTCAAA	AGTGTTATAA	GAAGATATTC	120
CTTTTTTTGT	CCTAGAGAAC	TTATTTTCCT	GTGAAA	ATG	CCT ACC ACA AAG AAG	174
				Met	Pro Thr Thr Lys Lys	
				1	5	
ACA TTG ATG TTC TTA TCA AGC TTT TTC ACC AGC CTT GGG TCC TTC ATT	222					
Thr Leu Met Phe Leu Ser Ser Phe Phe Thr Ser Leu Gly Ser Phe Ile						
10 15 20						
GTA ATT TGC TCT ATT CTT GGG ACA CAA GCA TGG ATC ACC AGT ACA ATT	270					
Val Ile Cys Ser Ile Leu Gly Thr Gln Ala Trp Ile Thr Ser Thr Ile						
25 30 35						
GCT GTT AGA GAC TCT GCT TCA AAT GGG AGC ATT TTC ATC ACT TAC GGA	318					
Ala Val Arg Asp Ser Ala Ser Asn Gly Ser Ile Phe Ile Thr Tyr Gly						
40 45 50						
CTT TTT CGT GGG GAG AGT AGT GAA GAA TTG AGT CAC GGA CTT GCA GAA	366					
Leu Phe Arg Gly Glu Ser Ser Glu Glu Leu Ser His Gly Leu Ala Glu						
55 60 65 70						
CCA AAG AAA AAG TTT GCA GTT TTA GAG ATA CTG AAT AAT TCT TCC CAA	414					
Pro Lys Lys Lys Phe Ala Val Leu Glu Ile Leu Asn Asn Ser Ser Gln						
75 80 85						
AAA ACT CTG CAT TCG GTG ACT ATC CTG TTC CTG GTC CTG AGT TTG ATC	462					
Lys Thr Leu His Ser Val Thr Ile Leu Phe Leu Val Leu Ser Leu Ile						
90 95 100						
ACG TCG CTG CTG AGC TCT GGG TTT ACC TTC TAC AAC AGC ATC AGC AAC	510					
Thr Ser Leu Leu Ser Ser Gly Phe Thr Phe Tyr Asn Ser Ile Ser Asn						
105 110 115						

151

CCT TAC CAG ACA TTC CTG GGG CCG ACG GGG GTG TAC ACC TGG AAC GGG 558  
 Pro Tyr Gln Thr Phe Leu Gly Pro Thr Gly Val Tyr Thr Trp Asn Gly  
 120 125 130  
 CTC GGT GCA TCC TTC GTT TTT GTG ACC ATG ATA CTG TTT GTG GCG AAC 606  
 5 Leu Gly Ala Ser Phe Val Phe Val Thr Met Ile Leu Phe Val Ala Asn  
 135 140 145 150  
 ACG CAG TCC AAC CAA CTC TCC GAA GAG TTG TTC CAA ATG CTT TAC CCG 654  
 Thr Gln Ser Asn Gln Leu Ser Glu Glu Leu Phe Gln Met Leu Tyr Pro  
 155 160 165  
 10 GCA ACC ACC AGT AAA GGA ACG ACC CAC AGT TAC GGA TAC TCG TTC TGG 702  
 Ala Thr Thr Ser Lys Gly Thr Thr His Ser Tyr Gly Tyr Ser Phe Trp  
 170 175 180  
 CTC ATA CTG CTC GTC ATT CTT CTA AAT ATA GTC ACT GTA ACC ATC ATC 750  
 Leu Ile Leu Leu Val Ile Leu Leu Asn Ile Val Thr Val Thr Ile Ile  
 15 185 190 195  
 ATT TTC TAC CAG AAG GCC AGA TAC CAG CGG AAG CAG GAG CAG AGA AAG 798  
 Ile Phe Tyr Gln Lys Ala Arg Tyr Gln Arg Lys Gln Glu Gln Arg Lys  
 200 205 210  
 CCA ATG GAA TAT GCT CCA AGG GAC GGA ATT TTA TTC TGAATTCTCT TTCATC 850  
 20 Pro Met Glu Tyr Ala Pro Arg Asp Gly Ile Leu Phe  
 215 220 225  
 TCATTTTGGC GTTGCATCTA TTGTACATCA GCCCTGAGTA GTAAGTGGTT AGCTTCTCTG 910  
 GACAATTACAG CATGGTAACG TGACTGTCAT CTGTGACAGC ATTTGTGTTT CATGACACTG 970  
 TGTTCTTCAT TGATGCTGTA CTCCTGAAAA TTTTCCAC AAGGTTGGGG AAATGAATGG 1030  
 25 GAAATGTGCG TGG 1043

## (2) INFORMATION FOR SEQ ID NO: 52:

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 972  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

35

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Liver

09445558 120199



(A) CHARACTERIZATION CODE: CDS

(C) CHARACTERIZATION METHOD: E

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

10	AGACAGCGGC GGGCGCAGGA CGTGCACT ATG GCT CGG GGC TCG CTG CGC CGG	52
	Met Ala Arg Gly Ser Leu Arg Arg	
	1 5	
	TTG CTG CGG CTC CTC GTG CTG GGG CTC TGG CTG GCG TTG CTG CGC TCC	100
	Leu Leu Arg Leu Leu Val Leu Gly Leu Trp Leu Ala Leu Leu Arg Ser	
15	10 15 20	
	GTG GCC GGG GAG CAA GCG CCA GGC ACC GCC CCC TGC TCC CGC GGC AGC	148
	Val Ala Gly Glu Gln Ala Pro Gly Thr Ala Pro Cys Ser Arg Gly Ser	
	25 30 35 40	
	TCC TGG AGC GCG GAC CTG GAC AAG TGC ATG GAC TGC GCG TCT TGC AGG	196
20	Ser Trp Ser Ala Asp Leu Asp Lys Cys Met Asp Cys Ala Ser Cys Arg	
	45 50 55	
	GCG CGA CCG CAC AGC GAC TTC TGC CTG GGC TGC GCT GCA GCA CCT CCT	244
	Ala Arg Pro His Ser Asp Phe Cys Leu Gly Cys Ala Ala Ala Pro Pro	
	60 65 70	
25	GCC CCC TTC CGG CTG CTT TGG CCC ATC CTT GGG GGC GCT CTG AGC CTG	292
	Ala Pro Phe Arg Leu Leu Trp Pro Ile Leu Gly Gly Ala Leu Ser Leu	
	75 80 85	
	ACC TTC GTG CTG GGG CTG CTT TCT GGC TTT TTG GTC TGG AGA CGA TGC	340
	Thr Phe Val Leu Gly Leu Leu Ser Gly Phe Leu Val Trp Arg Arg Cys	
30	90 95 100	
	CGC AGG AGA GAG AAG TTC ACC ACC CCC ATA GAG GAG ACC GGC GGA GAG	388
	Arg Arg Arg Glu Lys Phe Thr Thr Pro Ile Glu Glu Thr Gly Gly Glu	
	105 110 115 120	
	GGC TGC CCA GCT GTG GCG CTG ATC CAG TGACA ATGT GCCCCCTGCC A CCGG	440
35	Gly Cys Pro Ala Val Ala Leu Ile Gln	
	125	
	GGCTCGCCCA CTCATCATTC ATTCAATCCAT TCTAGAGCCA GTCTCTGCCT CCCAGACGCG	500
	GCGGGAGCCA AGCTCCTCCA ACCACAAGGG GGGTGGGGGG CGGTGAATCA CCTCTGAGGC	560

153

CTGGGCCCAG GGTTCAGGGG AACCTTCCAA GGTGTCTGGT TGCCCTGCCT CTGGCTCCAG 620  
 AACAGAAAGG GAGCCTCACG CTGGCTCACA CAAAACAGCT GACACTGACT AAGGAACTGC 680  
 AGCATTGCA CAGGGGAGGG GGGTGCCCTC CTTCTAGAG GCCCTGGGGG CCAGGCTGAC 740  
 TTGGGGGGCA GACTTGACAC TAGGCCCCAC TCACTCAGAT GTCCTGAAAT TCCACCACGG 800  
 5 GGGTCACCCT GGGGGGTTAG GGACCTATTT TTAACACTAG GGGGCTGGCC CACTAGGAGG 860  
 GCTGGCCCTA AGATACAGAC CCCCCCAACT CCCCAAAGCG GGGAGGAGAT ATTTATTTTG 920  
 GGGAGAGTTT GGAGGGGAGG GAGAATTTAT TAATAAAAGA ATCTTAACT TT 972

## 10 (2) INFORMATION FOR SEQ ID NO: 53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 695  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 15 (D) TOPOLOGY: Linear

## (ii) SEQUENCE KIND: cDNA to mRNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 20 (B) CELL KIND: Liver  
 (C) CELL LINE:  
 (D) CLONE NAME: HP10433

## (ix) SEQUENCE CHARACTERISTICS:

- 25 (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 73.. 564  
 (C) CHARACTERIZATION METHOD: E

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

30 AAGATTTTCAG CTGCGGGACG GTCAGGGGAG ACCTCCAGGC GCAGGGAAGG ACGGCCAGGG 60  
 TGACACGGAA GC ATG CGA CGG CTG CTG ATC CCT CTG GCC CTG TGG CTG GGC 111  
 Met Arg Arg Leu Leu Ile Pro Leu Ala Leu Trp Leu Gly  
 1 5 10  
 35 GCG GTG GGC GTG GGC GTC GCC GAG CTC ACG GAA GCC CAG CGC CGG GGC 159  
 Ala Val Gly Val Gly Val Ala Glu Leu Thr Glu Ala Gln Arg Arg Gly  
 15 20 25  
 CTG CAG GTG GCC CTG GAG GAA TTT CAC AAG CAC CCG CCC GTG CAG TGG 207

04453-1019  
66T02T-85254460

154

Leu Gln Val Ala Leu Glu Glu Phe His Lys His Pro Pro Val Gln Trp  
 30 35 40 45  
 GCC TTC CAG GAG ACC AGT GTG GAG AGC GCC GTG GAC ACG CCC TTC CCA 255  
 Ala Phe Gln Glu Thr Ser Val Glu Ser Ala Val Asp Thr Pro Phe Pro  
 5 50 55 60  
 GCT GGA ATA TTT GTG AGG CTG GAA TTT AAG CTG CAG CAG ACA AGC TGC 303  
 Ala Gly Ile Phe Val Arg Leu Glu Phe Lys Leu Gln Gln Thr Ser Cys  
 65 70 75  
 CGG AAG AGG GAC TGG AAG AAA CCC GAG TGC AAA GTC AGG CCC AAT GGG 351  
 10 Arg Lys Arg Asp Trp Lys Lys Pro Glu Cys Lys Val Arg Pro Asn Gly  
 80 85 90  
 AGG AAA CGG AAA TGC CTG GCC TGC ATC AAA CTG GGC TCT GAG GAC AAA 399  
 Arg Lys Arg Lys Cys Leu Ala Cys Ile Lys Leu Gly Ser Glu Asp Lys  
 95 100 105  
 15 GTT CTG GGC CGG TTG GTC CAC TGC CCC ATA GAG ACC CAA GTT CTG CGG 447  
 Val Leu Gly Arg Leu Val His Cys Pro Ile Glu Thr Gln Val Leu Arg  
 110 115 120 125  
 GAG GCT GAG GAG CAC CAG GAG ACC CAG TGC CTC AGG GTG CAG CGG GCT 495  
 Glu Ala Glu Glu His Gln Glu Thr Gln Cys Leu Arg Val Gln Arg Ala  
 20 130 135 140  
 GGT GAG GAC CCC CAC AGC TTC TAC TTC CCT GGA CAG TTC GCC TTC TCC 543  
 Gly Glu Asp Pro His Ser Phe Tyr Phe Pro Gly Gln Phe Ala Phe Ser  
 145 150 155  
 AAG GCC CTG CCC CGC AGC TAAGCCAGCA CTGAGCTGCG TGGTGCCTC 590  
 25 Lys Ala Leu Pro Arg Ser  
 160  
 CAGGACCGCT GCCGGTGGTA ACCAGTGGAA GACCCCAGCC CCCAGGGAGA GGACCCCGTT 650  
 CTATCCCCAG CCATGATAAT AAAGCTGCTC TCCCAGCTGC CTCTC 695

30

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1914

(B) TYPE: Nucleic acid

35

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) SEQUENCE KIND: cDNA to mRNA

044558-120196

155

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Homo sapiens*  
 (B) CELL KIND: Stomach cancer  
 (D) CLONE NAME: HP10480

5

## (ix) SEQUENCE CHARACTERISTICS:

- (A) CHARACTERIZATION CODE: CDS  
 (B) EXISTENCE POSITION: 80.. 661  
 (C) CHARACTERIZATION METHOD: E

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

ACTCTCTGCT GTCGCCCCGTC CCGCGCGCTC CTCCGACCCG CTCCGCTCCG CTCCGCTCGG 60  
 CCCC GCGCCG CCCGTCAAC ATG ATC CGC TGC GGC CTG GCC TGC GAG CGC TGC 112  
 Met Ile Arg Cys Gly Leu Ala Cys Glu Arg Cys  
 1 5 10  
 CGC TGG ATC CTG CCC CTG CTC CTA CTC AGC GCC ATC GCC TTC GAC ATC 160  
 Arg Trp Ile Leu Pro Leu Leu Leu Leu Ser Ala Ile Ala Phe Asp Ile  
 15 20 25  
 ATC GCG CTG GCC GGC CGC GGC TGG TTG CAG TCT AGC GAC CAC GGC CAG 208  
 Ile Ala Leu Ala Gly Arg Gly Trp Leu Gln Ser Ser Asp His Gly Gln  
 30 35 40  
 ACG TCC TCG CTG TGG TGG AAA TGC TCC CAA GAG GGC GGC GGC AGC GGG 256  
 Thr Ser Ser Leu Trp Trp Lys Cys Ser Gln Glu Gly Gly Gly Ser Gly  
 25 45 50 55  
 TCC TAC GAG GAG GGC TGT CAG AGC CTC ATG GAG TAC GCG TGG GGT AGA 304  
 Ser Tyr Glu Glu Gly Cys Gln Ser Leu Met Glu Tyr Ala Trp Gly Arg  
 60 65 70 75  
 GCA GCG GCT GCC ATG CTC TTC TGT GGC TTC ATC ATC CTG GTG ATC TGT 352  
 Ala Ala Ala Ala Met Leu Phe Cys Gly Phe Ile Ile Leu Val Ile Cys  
 30 80 85 90  
 TTC ATC CTC TCC TTC TTC GCC CTC TGT GGA CCC CAG ATG CTT GTC TTC 400  
 Phe Ile Leu Ser Phe Phe Ala Leu Cys Gly Pro Gln Met Leu Val Phe  
 95 100 105  
 CTG AGA GTG ATT GGA GGT CTC CTT GCC TTG GCT GCT GTG TTC CAG ATC 448  
 Leu Arg Val Ile Gly Gly Leu Leu Ala Leu Ala Ala Val Phe Gln Ile  
 110 115 120  
 ATC TCC CTG GTA ATT TAC CCC GTG AAG TAC ACC CAG ACC TTC ACC CTT 496

0944553 "12019" 85254460

156

Ile Ser Leu Val Ile Tyr Pro Val Lys Tyr Thr Gln Thr Phe Thr Leu  
125 130 135  
CAT GCC AAC CGT GCT GTC ACT TAC ATC TAT AAC TGG GCC TAC GGC TTT 544  
His Ala Asn Arg Ala Val Thr Tyr Ile Tyr Asn Trp Ala Tyr Gly Phe  
5 140 145 150 155  
GGG TGG GCA GCC ACG ATT ATC CTG ATC GGC TGT GCC TTC TTC TTC TGC 592  
Gly Trp Ala Ala Thr Ile Ile Leu Ile Gly Cys Ala Phe Phe Phe Cys  
160 165 170  
TGC CTC CCC AAC TAC GAA GAT GAC CTT CTG GGC AAT GCC AAG CCC AGG 640  
10 Cys Leu Pro Asn Tyr Glu Asp Asp Leu Leu Gly Asn Ala Lys Pro Arg  
175 180 185  
TAC TTC TAC ACA TCT GCC TA ACTTGGG AATGAATGTG GGAGAAAATC GCT 690  
Tyr Phe Tyr Thr Ser Ala  
190  
15 GCTGCTGAGA TGGACTCCAG AAGAAGAAAC TGTTTCTCCA GGCGACTTTG AACCCATTTT 750  
TTGGCAGTGT TCATATTATT AAAGTAGTCA AAAATGCTAA AATAATTTGG GAGAAAATAT 810  
TTTTTAAGTA GTGTTATAGT TTCATGTTTA TCTTTTATTA TGTTTTGTGA AGTTGTGTCT 870  
TTTCACTAAT TACCTATACT ATGCCAATAT TTCCTTATAT CTATCCATAA CATTTATACT 930  
ACATTTGTAA GAGAATATGC ACGTGAAACT TAACACTTTA TAAGGTAAAA ATGAGGTTTC 990  
20 CAAGATTTAA TAATCTGATC AAGTTCTTGT TATTTCCAAA TAGAATGGAC TTGGTCTGTT 1050  
AAGGGCTAAG GAGAAGAGGA AGATAAGGTT AAAAGTTGTT AATGACCAAA CATTCTAAAA 1110  
GAAATGCAAA AAAAAAGTTT ATTTTCAAGC CTTCGAACTA TTTAAGGAAA GCAAAATCAT 1170  
TTCCTAAATG CATATCATTT GTGAGAATTT CTCATTAATA TCCTGAATCA TTCATTTTCA 1230  
CTAAGGCTTC ATGTTGACTC GATATGTCAT CTAGGAAAGT ACTATTTTCAT GGTCCAAACC 1290  
25 TGTTGCCATA GTTGTAAGG CTTTCCTTTA AGTGTGAAAT ATTTAGATGA AATTTTCTCT 1350  
TTTAAAGTTC TTTATAGGGT TAGGGTGTGG GAAAATGCTA TATTAATAAA TCTGTAGTGT 1410  
TTTGTGTTTA TATGTTTACA ACCAGAGTAG ACTGGATTGA AAGATGGACT GGGTCTAATT 1470  
TATCATGACT GATAGATCTG GTTAAGTTGT GTAGTAAAGC ATTAGGAGGG TCATTCTTGT 1530  
CACAAAAGTG CCACTAAAAC AGCCTCAGGA GAATAAATGA CTTGCTTTTC TAAATCTCAG 1590  
30 GTTTATCTGG GCTCTATCAT ATAGACAGGC TTCTGATAGT TTGCAACTGT AAGCAGAAAC 1650  
CTACATATAG TTAAATCCT GGTCTTTCTT GGTAACAGA TTTTAAATGT CTGATATAAA 1710  
ACATGCCACA GGAGAATTCG GGGATTTGAG TTTCTCTGAA TAGCATATAT ATGATGCATC 1770  
GGATAGGTCA TTATGATTTT TTACCATTTT GACTTACATA ATGAAAACCA ATTCATTTTA 1830  
AATATCAGAT TATTATTTTG TAAGTTGTGG AAAAAGCTAA TTGTAGTTTT CATTATGAAG 1890  
35 TTTTCCAAT AAACCAGGTA TTCT 1914

044555 120199 06T02T 8224460

Sub  
a2  
cont

## PATENT COOPERATION TREATY

## PCT

REC'D 20 SEP 1999

WIPO PCT

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 660479	<b>FOR FURTHER ACTION</b> See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/JP98/02445	International filing date (day/month/year) 03/06/1998	Priority date (day/month/year) 03/06/1997
International Patent Classification (IPC) or national classification and IPC C12N15/12		
Applicant SAGAMI CHEMICAL RESEARCH CENTER et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.



2. This REPORT consists of a total of 5 sheets, including this cover sheet.

- ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☒ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand 17/12/1998	Date of completion of this report 14.09.99
Name and mailing address of the international preliminary examining authority:  European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Kaas, V Telephone No. +49 89 2399 8704 

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/JP98/02445

## I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

### Description, pages:

1-156 as originally filed

### Claims, No.:

1-6 as originally filed

### Drawings, sheets:

1/19-19/19 as originally filed

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

## III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

- ☐ the entire international application.
- ☒ claims Nos. 1-6 (all partially).

because:

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/JP98/02445

- ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):
  
- ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):
  
- ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.
  
- ☒ no international search report has been established for the said claims Nos. 1-6 (all partially).

**IV. Lack of unity of invention**

1. In response to the invitation to restrict or pay additional fees the applicant has:

- ☐ restricted the claims.
- ☐ paid additional fees.
- ☐ paid additional fees under protest.
- ☒ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- ☐ complied with.
- ☒ not complied with for the following reasons:

**see separate sheet**

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- ☐ all parts.
- ☒ the parts relating to claims Nos. 1-6 (all partially).



**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/JP98/02445

---

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

**1. Statement**

Novelty (N)	Yes:	Claims	1-6
	No:	Claims	
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-6
Industrial applicability (IA)	Yes:	Claims	1-6
	No:	Claims	

**2. Citations and explanations**

**see separate sheet**

1) For the reasons given in the International Preliminary Search Report, the application is considered to lack unity of invention (Rule 13 PCT). Given that no required additional search fees were timely paid by the applicant, the present IPER is limited to invention 1 (Claims 1-6, all partially) as mentioned therein.

2) Claims 1-4 relate to a cDNA clone (SEQ ID NO: 19 and 37) identified with the help of automated processing from a human liver cDNA library and to the polypeptide (SEQ ID NO: 1) allegedly expressed by said clone. However, the application fails to provide any evidence that the clone is capable of expressing a protein which is functional. In this respect, the mere fact that a protein appears to display a hydrophobic region of putative transmembrane domain(s) is not considered to represent such evidence. Moreover, there appears to be no substantial support therein for any of the numerous alleged putative technical useful properties which are merely listed in the description on pages 56 to 82. The present application is therefore considered to provide a human cDNA sequence encoding a polypeptide with no identified biological properties.

In this case, any prior art compound, regardless of its technical properties, is equally suitable as the starting point for making structural modifications and may be considered to represent the closest prior art.

Starting from this point, the objective technical problem to be solved can thus be formulated to lie in the provision of further human cDNA sequences as such, regardless of their possible useful properties. Without the concomitant need to provide any particular technical effect, the skilled person would have the choice of an infinite number of equally obvious possible solutions. In this respect, any DNA sequence of a human cDNA library is considered to obviously solve the above problem. The claimed polynucleotide molecule therefore represents the result of an arbitrary selection among the cDNA clones which can be routinely isolated from a human cDNA library, e.g. identified by querying an EST database for sequences possessing a putative secretion signal. Such an arbitrary selection cannot involve an inventive step because it is not justified by any technical purpose related to an hitherto unknown or unexpected technical effect resulting from those structural features which distinguish the compound claimed from all other possibilities.

Thus, for nucleotide and peptide sequences whose function and, in the latter case, existence is based purely upon surmise, inventive step cannot be acknowledged.

Claims 1-4 therefore do not satisfy the criteria as set forth in Article 33(3) PCT.

3) The preparation of a conventional expression vector comprising portions of the nucleotide sequence as well as the transformation of an eucaryotic cell with said vector do not appear to go beyond the knowledge and capabilities of the skilled person. Claims 5 and 6 therefore also lack inventiveness.

4) Claims 1-6 of invention 1 appear to be susceptible of industrial applicability as defined in Article 33(4) PCT.

## PATENT COOPERATION TREATY

PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

United States Patent and Trademark  
Office  
(Box PCT)  
Crystal Plaza 2  
Washington, DC 20231  
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year)

28 January 1999 (28.01.99)

International application No.

PCT/JP98/02445

Applicant's or agent's file reference

660479

International filing date (day/month/year)

03 June 1998 (03.06.98)

Priority date (day/month/year)

03 June 1997 (03.06.97)

Applicant

KATO, Seishi et al

1. The designated Office is hereby notified of its election made:



in the demand filed with the International Preliminary Examining Authority on:

17 December 1998 (17.12.98)



in a notice effecting later election filed with the International Bureau on:

2. The election ☒ was

was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

K. Takeda

Telephone No.: (41-22) 338.83.38

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference <b>660479</b>	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/JP 98/02445</b>	International filing date (day/month/year) <b>03/06/1998</b>	(Earliest) Priority Date (day/month/year) <b>03/06/1997</b>
Applicant <b>SAGAMI CHEMICAL RESEARCH CENTER et al.</b>		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (see Box I).

2. ☒ Unity of invention is lacking (see Box II).

3. ☒ The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing

☒ filed with the international application.

☐ furnished by the applicant separately from the international application,

☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.

☐ Transcribed by this Authority

4. With regard to the title, ☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is:

Figure No. \_\_\_\_\_ ☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☒ None of the figures.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 98/02445

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/705 A61K38/17 C12N5/10 C12Q1/37  
C12N9/72 C12N15/85

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KYTE J. ET AL.: "A SIMPLE METHOD FOR DISPLAYING THE HYDROPATHIC CHARACTER OF A PROTEIN" JOURNAL OF MOLECULAR BIOLOGY, vol. 157, no. 1, 5 May 1982, pages 105-132, XP000609692 cited in the application ---	
A	LIBERT F. ET AL.: "SELECTIVE AMPLIFICATION AND CLONING OF FOUR NEW MEMBERS OF THE G PROTEIN-COUPLED RECEPTOR FAMILY" SCIENCE, vol. 244, 5 May 1989, pages 569-572, XP002041588 --- -/--	

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## ° Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

22 September 1998

Date of mailing of the international search report

26. 01. 99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 98/02445

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MILLS A. AND DUGGAN M.J.: "ORPHAN SEVEN TRANSMEMBRANE DOMAIN RECEPTORS: REVERSING PHARMACOLOGY" TRENDS IN BIOTECHNOLOGY, vol. 12, February 1994, pages 47-49, XP002078287 ---	
A	Database EMBL, entry Emest7:HS010272 Accession number N39010 25 January 1996 99% identity with Seq.ID:19 nt.647-1146. XP002078288 see the whole document ---	2-4
A	Database EMBL, entry Emest9:HS204207 Accession number H57204 7 October 1995 96% identity with Seq.ID:19 nt.1-437. XP002078292 cited in the application see the whole document -----	2-4

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP 98/02445

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-6 : all partially (see subject 1, extra sheet)

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

1. Claims: 1-6 all partially.

A protein comprising an aminoacid sequence as in Seq.ID:1, encoding DNA, as in Seq.ID:19 and 37, related expression vector and transformed eukaryotic cell.

2. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:2, 20 and 38.

3. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:3, 21 and 39.

4. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:4, 22 and 40.

5. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:5, 23 and 41.

6. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:6, 24 and 42.

7. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:7, 25 and 43.

8. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:8, 26 and 44.

9. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:9, 27 and 45.

10. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:10, 28 and 46.

11. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:11, 29 and 47.



**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

12. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:12, 30 and 48.

13. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:13, 31 and 49.

14. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:14, 32 and 50.

15. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:15, 33 and 51.

16. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:16, 34 and 52.

17. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:17, 35 and 53.

18. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:18, 36 and 54.

## PATENT COOPERATION TREATY

PCT

From the INTERNATIONAL BUREAU

NOTIFICATION OF THE RECORDING  
OF A CHANGE(PCT Rule 92bis.1 and  
Administrative Instructions, Section 422)

To:

AOYAMA, Tamotsu  
Aoyama & Partners  
IMP Building  
3-7, Shiromi 1-chome  
Chuo-ku, Osaka-shi  
Osaka 540-0001  
JAPON

Date of mailing (day/month/year) 29 March 1999 (29.03.99)	IMPORTANT NOTIFICATION
Applicant's or agent's file reference 660479	
International application No. PCT/JP98/02445	International filing date (day/month/year) 03 June 1998 (03.06.98)

## 1. The following indications appeared on record concerning:

☒ the applicant      ☒ the inventor      ☐ the agent      ☐ the common representative

Name and Address YAMAGUCHI, Tomoko 5-13-11, Takasago Katsushika-ku Tokyo 125 Japan	State of Nationality JP	State of Residence JP
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

## 2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☐ the person      ☒ the name      ☒ the address      ☐ the nationality      ☐ the residence

Name and Address KIMURA, Tomoko 302, 4-1-28, Nishiikuta Tama-ku Kawasaki-shi Kanagawa 214-0037 Japan	State of Nationality JP	State of Residence JP
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

## 3. Further observations, if necessary:

## 4. A copy of this notification has been sent to:

<input checked="" type="checkbox"/> the receiving Office	<input type="checkbox"/> the designated Offices concerned
<input type="checkbox"/> the International Searching Authority	<input checked="" type="checkbox"/> the elected Offices concerned
<input checked="" type="checkbox"/> the International Preliminary Examining Authority	<input type="checkbox"/> other:

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland Facsimile No.: (41-22) 740.14.35	Authorized officer K. Takeda Telephone No.: (41-22) 338.83.38
---	---

# PATENT COOPERATION TREATY

From the  
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

AOYAMA, Tamotsu et al  
Aoyama & Partners  
IMP Building, 3-7 Shiromi 1-chome  
Chuo-ku, Osaka-shi  
Osaka 540  
JAPON



## PCT

### NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT (PCT Rule 71.1)

Date of mailing  
(day/month/year)

1 4. 09. 99

Applicant's or agent's file reference  
660479

#### IMPORTANT NOTIFICATION

International application No.  
PCT/JP98/02445

International filing date (day/month/year)  
03/06/1998

Priority date (day/month/year)  
03/06/1997

Applicant

SAGAMI CHEMICAL RESEARCH CENTER et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

#### 4. REMINDER

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Express Mail mailing label number: EE 622461032  
Date of Deposit: December 1, 1999  
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner For Patents, Washington, D.C. 20231

*[Signature]*

Name and mailing address of the IPEA/



European Patent Office  
D-80298 Munich  
Tel. +49 89 2399 - 0 Tx: 523656 epmu d  
Fax: +49 89 2399 - 4465

Authorized officer

Vullo, C

Tel. +49 89 2399-8061



# PATENT COOPERATION TREATY

## PCT

### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference <b>660479</b>	<b>FOR FURTHER ACTION</b>		See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. <b>PCT/JP98/02445</b>	International filing date (day/month/year) <b>03/06/1998</b>	Priority date (day/month/year) <b>03/06/1997</b>	
International Patent Classification (IPC) or national classification and IPC <b>C12N15/12</b>			
Applicant <b>SAGAMI CHEMICAL RESEARCH CENTER et al.</b>			

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.



2. This REPORT consists of a total of 5 sheets, including this cover sheet.

☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of    sheets.

3. This report contains indications relating to the following items:

- I    ☒ Basis of the report
- II   ☐ Priority
- III ☒ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☒ Lack of unity of invention
- V   ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☐ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand  <b>17/12/1998</b>	Date of completion of this report  <b>14. 09. 99</b>
Name and mailing address of the international preliminary examining authority:   <b>European Patent Office</b> <b>D-80298 Munich</b> <b>Tel. +49 89 2399 - 0 Tx: 523656 epmu d</b> <b>Fax: +49 89 2399 - 4465</b>	Authorized officer  <b>Kaas, V</b>  Telephone No. +49 89 2399 8704  

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/JP98/02445

## I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

### Description, pages:

1-156 as originally filed

### Claims, No.:

1-6 as originally filed

### Drawings, sheets:

1/19-19/19 as originally filed

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

## III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non-obvious), or to be industrially applicable have not been examined in respect of:

- ☐ the entire international application.
- ☒ claims Nos. 1-6 (all partially).

because:

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/JP98/02445

- ☐ the said international application, or the said claims Nos. relate to the following subject matter which does not require an international preliminary examination (*specify*):
  
- ☐ the description, claims or drawings (*indicate particular elements below*) or said claims Nos. are so unclear that no meaningful opinion could be formed (*specify*):
  
- ☐ the claims, or said claims Nos. are so inadequately supported by the description that no meaningful opinion could be formed.
- ☒ no international search report has been established for the said claims Nos. 1-6 (all partially).

**IV. Lack of unity of invention**

1. In response to the invitation to restrict or pay additional fees the applicant has:

- ☐ restricted the claims.
- ☐ paid additional fees.
- ☐ paid additional fees under protest.
- ☒ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

- ☐ complied with.
- ☒ not complied with for the following reasons:

**see separate sheet**

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

- ☐ all parts.
- ☒ the parts relating to claims Nos. 1-6 (all partially).

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/JP98/02445

---

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

**1. Statement**

Novelty (N)	Yes:	Claims	1-6
	No:	Claims	
Inventive step (IS)	Yes:	Claims	
	No:	Claims	1-6
Industrial applicability (IA)	Yes:	Claims	1-6
	No:	Claims	

**2. Citations and explanations**

**see separate sheet**

1) For the reasons given in the International Preliminary Search Report, the application is considered to lack unity of invention (Rule 13 PCT). Given that no required additional search fees were timely paid by the applicant, the present IPER is limited to invention 1 (Claims 1-6, all partially) as mentioned therein.

2) Claims 1-4 relate to a cDNA clone (SEQ ID NO: 19 and 37) identified with the help of automated processing from a human liver cDNA library and to the polypeptide (SEQ ID NO: 1) allegedly expressed by said clone. However, the application fails to provide any evidence that the clone is capable of expressing a protein which is functional. In this respect, the mere fact that a protein appears to display a hydrophobic region of putative transmembrane domain(s) is not considered to represent such evidence. Moreover, there appears to be no substantial support therein for any of the numerous alleged putative technical useful properties which are merely listed in the description on pages 56 to 82. The present application is therefore considered to provide a human cDNA sequence encoding a polypeptide with no identified biological properties.

In this case, any prior art compound, regardless of its technical properties, is equally suitable as the starting point for making structural modifications and may be considered to represent the closest prior art.

Starting from this point, the objective technical problem to be solved can thus be formulated to lie in the provision of further human cDNA sequences as such, regardless of their possible useful properties. Without the concomitant need to provide any particular technical effect, the skilled person would have the choice of an infinite number of equally obvious possible solutions. In this respect, any DNA sequence of a human cDNA library is considered to obviously solve the above problem. The claimed polynucleotide molecule therefore represents the result of an arbitrary selection among the cDNA clones which can be routinely isolated from a human cDNA library, e.g. identified by querying an EST database for sequences possessing a putative secretion signal. Such an arbitrary selection cannot involve an inventive step because it is not justified by any technical purpose related to an hitherto unknown or unexpected technical effect resulting from those structural features which distinguish the compound claimed from all other possibilities.

Thus, for nucleotide and peptide sequences whose function and, in the latter case, existence is based purely upon surmise, inventive step cannot be acknowledged.

Claims 1-4 therefore do not satisfy the criteria as set forth in Article 33(3) PCT.

3) The preparation of a conventional expression vector comprising portions of the nucleotide sequence as well as the transformation of an eucaryotic cell with said vector do not appear to go beyond the knowledge and capabilities of the skilled person. Claims 5 and 6 therefore also lack inventiveness.

4) Claims 1-6 of invention 1 appear to be susceptible of industrial applicability as defined in Article 33(4) PCT.



From the INTERNATIONAL SEARCHING AUTHORITY

PCT

To:

Aoyama & Partners  
IMP Building, 3-7 Shiromi 1-chome  
Chuo-ku, Osaka-shi  
Osaka 540  
JAPAN

NOTIFICATION OF TRANSMITTAL OF  
THE INTERNATIONAL SEARCH REPORT  
OR THE DECLARATION



(PCT Rule 44.1)

Date of mailing  
(day/month/year)

26/01/1999

Applicant's or agent's file reference

660479

FOR FURTHER ACTION

See paragraphs 1 and 4 below

International application No.

PCT/JP 98/02445

International filing date  
(day/month/year)

03/06/1998

Applicant

SAGAMI CHEMICAL RESEARCH CENTER et al.

1. ☒ The applicant is hereby notified that the International Search Report has been established and is transmitted.

**Filing of amendments and statement under Article 19:**

The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

**When?** The time limit for filing such amendments is normally 2 months from the date of transmittal of the International Search Report; however, for more details, see the notes on the accompanying sheet.

**Where?** Directly to the International Bureau of WIPO  
34, chemin des Colombettes  
1211 Geneva 20, Switzerland  
Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no International Search Report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90 bis.1 and 90 bis.3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within 20 months from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

"Express Mail" mailing label number: 986320464494  
 Date of Deposit: December 1, 1999  
 I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner For Patents, Washington, D.C. 20231

Name and mailing address of the International Searching Authority



European Patent Office, P.B. 5818 Patentlaan 2  
NL-2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Andria Overbeeke-Siepkens

These Notes are intended to give the basic instructions concerning the filing of amendments under article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the PCT Applicant's Guide, a publication of WIPO.

In these Notes, "Article", "Rule", and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions respectively.

## INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only.

### What parts of the international application may be amended?

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

### When?

Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

### Where not to file the amendments?

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

### How?

Either by cancelling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Administrative Instructions, Section 205(b)).

The amendments must be made in the language in which the international application is to be published.

### What documents must/may accompany the amendments?

#### Letter (Section 205(b)):

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.

The letter must indicate the differences between the claims as filed and the claims as amended. It must, in particular, indicate, in connection with each claim appearing in the international application (it being understood that identical indications concerning several claims may be grouped), whether

- (i) the claim is unchanged;
- (ii) the claim is cancelled;
- (iii) the claim is new;
- (iv) the claim replaces one or more claims as filed;
- (v) the claim is the result of the division of a claim as filed.

**The following examples illustrate the manner in which amendments must be explained in the accompanying letter:**

1. [Where originally there were 48 claims and after amendment of some claims there are 51]:  
"Claims 1 to 29, 31, 32, 34, 35, 37 to 48 replaced by amended claims bearing the same numbers; claims 30, 33 and 36 unchanged; new claims 49 to 51 added."
2. [Where originally there were 15 claims and after amendment of all claims there are 11]:  
"Claims 1 to 15 replaced by amended claims 1 to 11."
3. [Where originally there were 14 claims and the amendments consist in cancelling some claims and in adding new claims]:  
"Claims 1 to 6 and 14 unchanged; claims 7 to 13 cancelled; new claims 15, 16 and 17 added." or  
"Claims 7 to 13 cancelled; new claims 15, 16 and 17 added; all other claims unchanged."
4. [Where various kinds of amendments are made]:  
"Claims 1-10 unchanged; claims 11 to 13, 18 and 19 cancelled; claims 14, 15 and 16 replaced by amended claim 14; claim 17 subdivided into amended claims 15, 16 and 17; new claims 20 and 21 added."

**"Statement under article 19(1)" (Rule 46.4)**

The amendments may be accompanied by a statement explaining the amendments and indicating any impact that such amendments might have on the description and the drawings (which cannot be amended under Article 19(1)).

The statement will be published with the international application and the amended claims.

**It must be in the language in which the international application is to be published.**

It must be brief, not exceeding 500 words if in English or if translated into English.

It should not be confused with and does not replace the letter indicating the differences between the claims as filed and as amended. It must be filed on a separate sheet and must be identified as such by a heading, preferably by using the words "Statement under Article 19(1)."

It may not contain any disparaging comments on the international search report or the relevance of citations contained in that report. Reference to citations, relevant to a given claim, contained in the international search report may be made only in connection with an amendment of that claim.

**Consequence if a demand for international preliminary examination has already been filed**

If, at the time of filing any amendments under Article 19, a demand for international preliminary examination has already been submitted, the applicant must preferably, at the same time of filing the amendments with the International Bureau, also file a copy of such amendments with the International Preliminary Examining Authority (see Rule 62.2(a), first sentence).

**Consequence with regard to translation of the international application for entry into the national phase**

The applicant's attention is drawn to the fact that, where upon entry into the national phase, a translation of the claims as amended under Article 19 may have to be furnished to the designated/elected Offices, instead of, or in addition to, the translation of the claims as filed.

For further details on the requirements of each designated/elected Office, see Volume II of the PCT Applicant's Guide.

## PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference <b>660479</b>	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/JP 98/ 02445</b>	International filing date (day/month/year) <b>03/06/1998</b>	(Earliest) Priority Date (day/month/year) <b>03/06/1997</b>
Applicant <b>SAGAMI CHEMICAL RESEARCH CENTER et al.</b>		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (see Box I).
2. ☒ **Unity of invention is lacking** (see Box II).
3. ☒ The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing
  - ☒ filed with the international application.
  - ☐ furnished by the applicant separately from the international application,
    - ☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.
  - ☐ Transcribed by this Authority
4. With regard to the title, ☒ the text is approved as submitted by the applicant.  
☐ the text has been established by this Authority to read as follows:
5. With regard to the abstract,
  - ☒ the text is approved as submitted by the applicant.
  - ☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.
6. The figure of the drawings to be published with the abstract is:
  - Figure No. \_\_\_\_\_ ☐ as suggested by the applicant.
  - ☐ because the applicant failed to suggest a figure.
  - ☐ because this figure better characterizes the invention.
  - ☒ None of the figures.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 98/02445

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/705 A61K38/17 C12N5/10 C12Q1/37  
C12N9/72 C12N15/85

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KYTE J. ET AL.: "A SIMPLE METHOD FOR DISPLAYING THE HYDROPATHIC CHARACTER OF A PROTEIN" JOURNAL OF MOLECULAR BIOLOGY, vol. 157, no. 1, 5 May 1982, pages 105-132, XP000609692 cited in the application ---	
A	LIBERT F. ET AL.: "SELECTIVE AMPLIFICATION AND CLONING OF FOUR NEW MEMBERS OF THE G PROTEIN-COUPLED RECEPTOR FAMILY" SCIENCE, vol. 244, 5 May 1989, pages 569-572, XP002041588 --- -/-	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*&\* document member of the same patent family

Date of the actual completion of the international search

22 September 1998

Date of mailing of the international search report

26. 01. 99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

## INTERNATIONAL SEARCH REPORT

International Application No

CT/JP 98/02445

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MILLS A. AND DUGGAN M.J.: "ORPHAN SEVEN TRANSMEMBRANE DOMAIN RECEPTORS: REVERSING PHARMACOLOGY" TRENDS IN BIOTECHNOLOGY, vol. 12, February 1994, pages 47-49, XP002078287 ---	
A	Database EMBL, entry Emest7:HS010272 Accession number N39010 25 January 1996 99% identity with Seq.ID:19 nt.647-1146. XP002078288 see the whole document ---	2-4
A	Database EMBL, entry Emest9:HS204207 Accession number H57204 7 October 1995 96% identity with Seq.ID:19 nt.1-437. XP002078292 cited in the application see the whole document -----	2-4

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP 98/02445

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-6 : all partially (see subject 1, extra sheet)

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

1. Claims: 1-6 all partially.

A protein comprising an aminoacid sequence as in Seq.ID:1, encoding DNA, as in Seq.ID:19 and 37, related expression vector and transformed eukaryotic cell.

2. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:2, 20 and 38.

3. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:3, 21 and 39.

4. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:4, 22 and 40.

5. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:5, 23 and 41.

6. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:6, 24 and 42.

7. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:7, 25 and 43.

8. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:8, 26 and 44.

9. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:9, 27 and 45.

10. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:10, 28 and 46.

11. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:11, 29 and 47.



**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

12. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:12, 30 and 48.

13. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:13, 31 and 49.

14. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:14, 32 and 50.

15. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:15, 33 and 51.

16. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:16, 34 and 52.

17. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:17, 35 and 53.

18. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:18, 36 and 54.

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference <b>660479</b>	<b>FOR FURTHER ACTION</b> see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. <b>PCT/JP 98/ 02445</b>	International filing date (day/month/year) <b>03/06/1998</b>	(Earliest) Priority Date (day/month/year) <b>03/06/1997</b>
Applicant <b>SAGAMI CHEMICAL RESEARCH CENTER et al.</b>		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (see Box I).

2. ☒ Unity of invention is lacking (see Box II).

3. ☒ The international application contains disclosure of a nucleotide and/or amino acid sequence listing and the international search was carried out on the basis of the sequence listing

☒ filed with the international application.

☐ furnished by the applicant separately from the international application,

☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.

☐ Transcribed by this Authority

4. With regard to the title, ☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the abstract,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.

6. The figure of the drawings to be published with the abstract is:

Figure No. \_\_\_\_\_ ☐ as suggested by the applicant.

☐ because the applicant failed to suggest a figure.

☐ because this figure better characterizes the invention.

☒ None of the figures.

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 98/02445

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/705 A61K38/17 C12N5/10 C12Q1/37  
C12N9/72 C12N15/85

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KYTE J. ET AL.: "A SIMPLE METHOD FOR DISPLAYING THE HYDROPATHIC CHARACTER OF A PROTEIN" JOURNAL OF MOLECULAR BIOLOGY, vol. 157, no. 1, 5 May 1982, pages 105-132, XP000609692 cited in the application	
A	LIBERT F. ET AL.: "SELECTIVE AMPLIFICATION AND CLONING OF FOUR NEW MEMBERS OF THE G PROTEIN-COUPLED RECEPTOR FAMILY" SCIENCE, vol. 244, 5 May 1989, pages 569-572, XP002041588 --- -/-	

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## ° Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

22 September 1998

Date of mailing of the international search report

26.01.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Macchia, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 98/02445

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MILLS A. AND DUGGAN M.J.: "ORPHAN SEVEN TRANSMEMBRANE DOMAIN RECEPTORS: REVERSING PHARMACOLOGY" TRENDS IN BIOTECHNOLOGY, vol. 12, February 1994, pages 47-49, XP002078287	
A	--- Database EMBL, entry Emest7:HS010272 Accession number N39010 25 January 1996 99% identity with Seq.ID:19 nt.647-1146. XP002078288 see the whole document	2-4
A	--- Database EMBL, entry Emest9:HS204207 Accession number H57204 7 October 1995 96% identity with Seq.ID:19 nt.1-437. XP002078292 cited in the application see the whole document -----	2-4

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/JP 98/02445

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-6 : all partially (see subject 1, extra sheet)

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

1. Claims: 1-6 all partially.

A protein comprising an aminoacid sequence as in Seq.ID:1, encoding DNA, as in Seq.ID:19 and 37, related expression vector and transformed eukaryotic cell.

2. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:2, 20 and 38.

3. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:3, 21 and 39.

4. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:4, 22 and 40.

5. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:5, 23 and 41.

6. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:6, 24 and 42.

7. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:7, 25 and 43.

8. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:8, 26 and 44.

9. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:9, 27 and 45.

10. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:10, 28 and 46.

11. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:11, 29 and 47.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

12. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:12, 30 and 48.

13. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:13, 31 and 49.

14. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:14, 32 and 50.

15. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:15, 33 and 51.

16. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:16, 34 and 52.

17. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:17, 35 and 53.

18. Claims: 1-6 all partially.

As invention 1 but concerning Seq.ID:18, 36 and 54.